RAINER BAUMANN

FAKTOREN PÄDAGOGISCHER PERSÖNLICHKEITSBEURTEILUNGEN

EINE UNTERSUCHUNG ZU KONSTANZ UND VERÄNDERUNG DER FAKTOREN DER IMPLIZITEN PERSÖNLICHKEITSTHEORIEN

Dissertation
zur
Erlangung des akademischen Grades
Doktor der Sozialwissenschaft
in der Fakultät
für Sozial- und Verhaltenswissenschaften
der Eberhard-Karls-Universität Tübingen

1980

1982

Rainer Baumann

Faktoren pädagogischer Persönlichkeitsbeurteilungen

Eine Untersuchung zu Konstanz und Veränderung der Faktoren der impliziten Persönlichkeitstheorien

Dissertation
zur
Erlangung des akademischen Grades
Doktor der Sozialwissenschaft
in der Fakultät
für Sozial- und Verhaltenswissenschaften
der Eberhard-Karls-Universität Tübingen

1982

Gedruckt mit Genehmigung
der Fakultät für Sozial- und Verhaltenswissenschaften
der Universität Tübingen

Hauptberichterstatter: Prof. Dr. Gerhard Kaminski
Mitberichterstatter: Prof. Dr. Heinz Mandl
Dekan: Prof. Dr. Gerhard Kaminski
Tag der mündlichen Prüfung: 1.Juli 1980

UNIV.
LEIDEN
BIBL.

Druck: Lutz Bongen , Moers

Vorwort

In vielen Redewendungen der Alltagssprache, die sich auf naive Persönlichkeitsdiagnostik beziehen, manifestiert sich das Phänomen des regelhaften gemeinsamen Auftretens von Eigenschaftsbezeichnungen. Wenn man auch der Validität dieser Ausflüsse des "gesunden Menschenverstandes" oder der "Volksweisheit" seitens der Fachwissenschaften zurecht mißtraut, so sind doch diese Urteilstendenzen für ihre Benutzer deswegen nicht weniger von starken Evidenzgefühlen begleitet.

Wenn solche Tendenzen im Alltagsleben auch oft zur Tradierung von Mißverständnissen beitragen, so sind sie doch deswegen etwas weniger ernst zu nehmen, weil sie gewissermaßen nur Privatleute betreffen und die resultierenden Urteile von begrenzter Reichweite sind.
Diese Beurteilungstendenzen, die als "naive" oder "implizite" Persönlichkeitstheorien Eingang in die Literatur über die Urteilsbildung fanden und als Korrelationen von Eigenschaften in Beurteilungen definiert wurden, verlieren ihre Harmlosigkeit, sobald sie z.B. in der offiziellen pädagogischen Persönlichkeitsdiagnostik im Rahmen der Schule wirken. Dort sind nämlich Persönlichkeitsbeurteilungen mit weitreichenden Folgen für die Beurteilten verbunden, so daß höchstmögliche Gültigkeit angestrebt werden sollte. Die Validität der Urteile wird aber allein schon durch Beurteilungstendenzen beeinträchtigt, die im jeweiligen Beurteiler liegen, denn der Urteiler bringt in seinem Urteil nur seine ganz persönlichen Annahmen bezüglich der Persönlichkeitsstruktur seines sozialen Gegenübers zum Ausdruck; er fällt ein Urteil über den Schüler, in dem über einen hohen Prozentsatz der beurteilten Persönlichkeitsmerkmale gar keine Beobachtungen oder sonstige Informationen vorliegen außer denen, die ihm seine IPT bezüglich des Zusammenhanges der nicht beobachteten mit den beobachteten Merkmalen liefert. Die Urteile über Merkmale, zu denen keine Informationen vorliegen, werden dem Schüler dann attribuiert.

Da Urteile unter anderem auch Erwartungen gegenüber dem Beurteilten fundieren, ist eine gemäß dem Urteil erfolgende Selektion zukünftiger Wahrnehmungen sehr wahrscheinlich, d.h. der Urteiler wählt Informationen nach Kongruenz mit seinem Urteil aus.

Zudem sorgt die Relevanz des Urteils für das aktuelle Verhalten des urteilenden Lehrers dafür, daß dem Schüler die Meinung des Lehrers über ihn transparent wird; der Schüler reagiert seinerseits so auf das Verhalten des Lehrers, daß sich ein Effekt der "self-fulfilling prophecies" (ROSENTHAL,R., JACOBSEN,L., 1968) einstellt, womit er sich in seinem Verhalten dem Urteil des Lehrers annähert und dieses bestätigt.

Nun ist ein Kreislauf geschlossen, der, um ein vielzitiertes Wort zu gebrauchen, oft genug aus einem Urteil eine Verurteilung werden läßt.
Wenn in der Folge oft von der 'impliziten Persönlichkeitstheorie' im Singular die Rede ist, so ist dies nicht durch die Vorstellung des Untersuchers bedingt, es gebe nur eine einzige oder jeder Urteiler verfüge nur über eine solche Theorie, die es festzustellen gilt.
Die impliziten Persönlichkeitstheorien werden vielmehr in dieser Arbeit als gruppenstereotype Merkmalskorrelationen in Urteilen angesehen, die wiederum von Urteiler zu Urteiler variieren können. Der Singular rührt von der zwangsweisen Mittelwertbildung, die mit der Untersuchung der Urteile von Gruppen über Gruppen verbunden ist. Dadurch können nur die kongruenten Momente aller Persönlichkeitstheorien der Beurteiler über alle Beurteilten erfaßt werden, was zu der obenerwähnten Formulierung führt.

Zuletzt sei der durch präzise Laborexperimente verwöhnte Leser noch um Nachsicht wegen mancher Insuffizienzen in Methode und statistischer Datenverarbeitung gebeten; der Untersucher hatte sich auf Neuland begeben, für das es zum Zeitpunkt der Konzipierung der Arbeit keine geeigneten Vorbildarbeiten gab, aus denen man sich hätte einen Rat holen können.

Außerdem ist die mit dieser Arbeit vorliegende Kombination von Feldstudie und Längsschnitt-Design von allen, die mit anderen Untersuchungsgegenständen Ähnliches gewagt hatten, als sehr dornenreicher Weg beschrieben worden, den man lieber zugunsten einer traditionelleren Konstellation aufgeben sollte.
Schwierigkeiten wie : Aufrechterhaltung der Motivation der Versuchspersonen, geringes Instrumentarium statistischer Verrechnungsmethoden, riesige Datenmengen bei gleichzeitig begrenzter Kapazität des verfügbaren Speicherplatzes in den Rechenzentren, sodaß die Möglichkeiten oftmals Methode und Ziel der Untersuchung eher bestimmten als der enthusiastische Plan des Untersuchers. Obwohl diese Effekte die Arbeit an dieser Studie oft stark beeinträchtigten, konnte hoffentlich ein gewisser methodischer und Ziel-Anspruch gewahrt werden.

Nun möchte ich besonders Herrn Prof.Dr.G.Kaminski danken für seine fachliche Beratung und seine unendliche Geduld, mit der er die vielen Verzögerungen tolerierte, die sich aus der ununterbrochenen Berufstätigkeit des Autors und der großen räumlichen Entfernung zur Universität ergaben.
Für ihren fachlichen Rat danken möchte ich auch Herrn Prof.Dr.H.Mandl vom Deutschen Institut für Fernstudien und Herrn Prof.Dr.E.Fürntratt vom Psychologischen Seminar der Pädagogischen Hochschule Rheinland, Abt. Aachen.

Aachen, September 1979 Rainer Baumann

Inhaltsverzeichnis

I Einführung

1. Beurteilung im Bereich der Schule 1
2. Naive Persönlichkeitstheorien von Lehrern 9

II Problemkreise mit Bezug zur 'impliziten Persönlichkeitstheorie'

2.1. Das Problem der Eigenschaften 12
2.2. Die soziale Wahrnehmung 18
2.3. Self-fulfilling prophecies 21
2.4. Selektivität der Wahrnehmung 26
2.5. Implizite Persönlichkeitstheorien 29
2.5.1. Theoretische Einordnung 29
2.5.2. Der Begriff "Implizite Persönlichkeitstheorie" 34
2.5.3. Interpretationsmodi der Impliziten Persönlichkeitstheorien 39
2.5.4. Methoden zur empirischen Erfassung von Persönlichkeitstheorien 41

III Fragestellung

3.1. Begleitumstände und Ziele der Arbeit und ihre Auswirkungen auf Fragestellung und Methodik 44
3.2. Vorbildarbeit 46
3.3. Eingrenzung der Fragestellung 49

IV Methode

4.1. Theoretische Fundierung der Variablenerfassung 57
4.2. Auswahl der Beurteiler-Stichprobe 58
4.2.1. Probleme der Stichprobe 59
4.2.2. Erfasste Variablen der Stichprobe 61
4.3. Auswahl der Beurteilten-Stichprobe 63
4.3.1. Erfasste Variablen der Stichprobe 64

4.4.	Methode zur Erfassung der impliziten Persönlichkeitstheorie durch Beurteilung von Stereotypen	65
4.4.1.	Das Meßinstrument	66
4.4.1.1.	Zwei typische Modelle in der 'person-perception'-Forschung	66
4.4.1.2.	Erhebung relevanter Eigenschaftsbegriffe	69
4.4.1.3.	Herstellung der Rating-Skalen	77
4.4.1.3.1.	Kennzeichnung der Pole	77
4.4.1.3.2.	Skalierung	79
4.4.1.3.3.	Auswahl der Stereotypen	80
4.4.1.3.4.	Statistische Bearbeitung der Daten	81
4.5.	Methode zur Erfassung der Faktorenstruktur aktueller Schülerbeurteilungen: Der Längsschnitt	82
4.5.1.	Probleme der Datensammlung in Längsschnitt-untersuchungen	82
4.5.1.1.	Design-Modelle	
4.5.1.2.	Probleme der Stichprobe	84
4.5.1.3.	Folgen wiederholter Datenerhebung	86
4.5.2.	Die Urteiler-Stichprobe (Lehrer)	88
4.5.2.2.	Die Beurteilten-Stichprobe (Schüler)	89
4.5.3.	Das Meßinstrument	90
4.5.3.1.	Die Rating-Liste	90
4.5.3.2.	Skala der "subjektiven Sicherheit"	92
4.5.4.	Versuchsplan I: Vergleich der IPT aus Stereotypenbeurteilungen mit den Faktoren-strukturen aktueller Schülerbeurteilungen	94
4.5.5.	Versuchsplan II - der Längsschnitt - Vergleich der Faktorenstrukturen aus den aktuellen Schülerbeurteilungen	96
4.5.6.	Statistische Methoden	98
4.5.6.1.	Probleme der Veränderungsmessung	99
4.5.6.2.	Validität und Reliabilität der Ergebnisse	106
4.5.6.3.	Statistische Methoden der Untersuchung	115
4.5.6.3.1.	Extraktion der Faktoren	116
4.5.6.3.2.	Sicherung der extrahierten Faktoren	117
4.5.6.3.3.	Identifikation der extrahierten Faktoren	118
4.5.6.3.4.	Interpretation der extrahierten Faktoren	119

V Ergebnisse

5.1.	Stichprobenbeschreibung	121
5.1.1.	Die Stichprobe der Beurteiler	121
5.1.1.1.	Zusammensetzung der Stichprobe	121
5.1.2.	Die Stichprobe der Beurteilten	124
5.2.	Faktorenanalyse der Stereotypen-beurteilungen	125
5.2.1.	Die Faktorenstruktur der Stereo-typenbeurteilungen	125
5.2.2.	Interpretation der Faktoren	127
5.3.	Faktorenanalyse aktueller Urteile über "unbekannte Schüler"	130
5.3.1.	Faktorenstruktur der Urteile	130
5.3.2.	Interpretation der Faktoren	132
5.4.	Kongruenz der Faktoren aus Stereo-typenbeurteilungen und aktuellen Beurteilungen relativ unbekannter Schüler	142
5.5.	Kongruenzen der Faktoren aus aktuellen Schülerbeurteilungen mit steigendem Bekanntschaftsgrad	145
5.5.1.	Die Faktorenlösungen	146
5.5.1.1.	Faktorenlösung nach dem Eigenwert-kriterium	147
5.5.1.1.1.	Faktorenextraktion	147
5.5.1.1.2.	Faktorenidentifikation	162
5.5.1.2.	Die 4-Faktoren-Lösung	174
5.5.1.2.1.	Faktorenextraktion	174
5.5.1.2.2.	Faktorenidentifikation	186
5.5.1.2.3.	Zusammenfassung der 4-Faktoren-Lösung	192
5.5.1.2.3.1.	Die zentrale Dimension	193
5.5.1.2.3.2.	Die zentrale Eigenschaft	195
5.6.	Einzeluntersuchungen über weitere Kennwerte der IPT	196
5.6.1.	Die Stereotypen	198
5.6.1.1.	Stereotyp: "der schlechte Schüler" der 1.Gymnasialklasse	199
5.6.1.2.	Stereotyp: "der gute Schüler" der 1.Gymnasialklasse	204
5.6.1.3.	Stereotyp: "der Schüler" der 1. Gymnasialklasse	207

5.6.2. Stereotypen und Lehrermerkmale 210
5.7. Einfluß der subjektiven Urteilssicherheit 213
5.8. Modifikation der Beurteilung über den Meßzeitraum 216
5.9. Stabilität der Persönlichkeitsbeurteilungen über den Meßzeitraum 227
5.10. Differenziertheit der Urteile 235

VI Ergebnisse 237

VII Zusammenfassung 255

Bibliographie 261

Anhang 280

I Einführung

1. Beurteilung im Bereich der Schule

Zwei der Hauptaufgaben, die von der Institution Schule an Lehrer herangetragen werden, sind die Vermittlung von Lerninhalten und die Überprüfung des fortschreitenden Erwerbs dieser Lerninhalte.

Obwohl schätzungsweise ein Viertel bis ein Drittel der späteren Funktionen der Lehrer unmittelbar auf Beurteilungen abzielen (KLEBER, 1976, S.20), existiert doch relativ zu diesem Verhältnis der späteren Funktionen, in der Ausbildung der Lehrer ein starkes Ungleichgewicht zugunsten der Behandlung des Vermittlungsaspekts. Es hat sich zwar eine umfangreiche Literatur zum Problemkreis des Beurteilens etabliert, diese wird jedoch nur in weit unterrepräsentativem Maße an die Lehramtsstudenten herangetragen:

> Bei der Durchsicht nur eines Vorlesungsverzeichnisses (WS 77/78) der Pädagogischen Hochschule Rheinland mit den Abteilungen Aachen, Bonn, Köln und Neuss wurden von den schulpädagogischen und psychologischen Seminaren insgesamt 315 Veranstaltungen für rund 15.000 Studierende angeboten. Davon hatten 96 Veranstaltungen explizit allgemein- oder fachdidaktischen Charakter, jedoch nur 7 = 2,2% befassten sich mit schulischer Leistungsdiagnostik.
> Im gleichen Zeitraum wurde im gleichen Einzugsbereich an den Ausbildungsstätten für Gymnasiallehrer (TH Aachen, Universitäten Köln und Bonn) keine Veranstaltung zu diesem Themenkreis durchgeführt.
> Auch in übergreifenderen erziehungswissenschaftlichen Aus- und Weiterbildungsmaßnahmen scheint diese Berufsfunktion nur gering reflektiert zu werden: Das sonst recht breit angelegte Funkkolleg "Erziehungswissenschaft" beschäftigt sich auf einer halben von 1000 im Jahr 1970 veröffentlichten Textseiten mit der Beurteilungsfunktion des Lehrers. (KLAFKI, W.et.al. (Hrsg.): Erziehungswissenschaft 1, 1973, 7, S.223).

Der geringen Vorbereitung der Lehrer auf ihre Beurteilungsaufgaben steht die pädagogische Bedeutung der Leistungsbeurteilung, die zumeist durch Notengebung erfolgt, gegenüber:

KLEBER (1976, S. 27 ff) legt zwei Funktionskreise der Beurteilung durch Notengebung dar, die wiederum in mehrere Einzelfunktionen zerfallen:

1. Funktionskreis: "Beurteilung zur Förderung" als "primär pädagogisches Anliegen"

Funktionen: 1.1. "Schüler-Feed-Back"

Der Schüler erfährt über die Beurteilung, wie weit er mit seinem derzeitigen Informationsniveau von dem als Lehrziel angestrebten entfernt ist. Die Beurteilung soll sein derzeitiges Lernverhalten bestärken oder verbessern.

1.2. "Lehrer-Feed-Back"

Der Lehrer erhält Rückmeldung über das Lernleistungsniveau seiner Schüler; er kann daraus Rückschlüsse über die Effektivität seines methodischen Vorgehens ziehen und dieses verbessern und differenzieren.

1.3. "Extrinsische Motivation des Schülers"

Unabhängig von sachorientierter Motivation kann der Schüler bei ergänzenden Maßnahmen des Elternhauses durch die Notengebung zu erhöhter Anstrengung motiviert werden.

1.4. "Extrinsische Motivation des Lehrers"

Der Lehrer kann bemüht sein, sein Lehrverhalten in Fächern und Klassen mit ungünstigem Notendurchschnitt zu verbessern, um seinem Arbeitgeber gegenüber zu belegen, daß er sein Gehalt zu Recht bezieht.

1.5. "Sozialisierung"

Die Schüler werden durch die Benotung mit Leistungsnormen als gesellschaftlich allgemein verbindlichen Standards vertraut gemacht im Gegensatz zu dem in der Familie geltenden individuellen Standards.

1.6. "Chancenausgleich"

Zur Förderung besonders benachteiligter Schüler kann eine zeitweise bewußt manipulierte Benotung vorgenommen werden zum Zweck der Motivierung.

1.7. "Hilfe für Plazierungsentscheidungen"

Stehen für das Erreichen eines Lehrzieles verschiedene methodische Wege zur Verfügung, kann die Beurteilung durch Notengebung helfen, dem Schüler den richtigen Lehrgang zuzuweisen.

2. Funktionskreis: "Auslesefunktion"

> Über die Leistungsbeurteilung erfolgt eine Gruppenzuteilung der Schüler, die Selektions- und Förderungsmaßnahmen dienen können. Ferner wird die Schullaufbahn eines Schülers durch die sich in den Benotungen widerspiegelnden Leistungen bestimmt.
> "Gleichzeitig verteilt über diese Auslese die Schule mit ihrer Benotung Berechtigung für die Vergabe von mehr oder weniger qualifizierten Tätigkeiten und sie dirigiert damit den sozialen Aufstieg." (KLEBER, 1976, S.30).

Weiter führt KLEBER (a.a.O.S.30 ff) noch die Berichtsfunktion der Benotung als Nachweis verantwortungsvoller Tätigkeit gegenüber Dienstherrn und Eltern auf und die Disziplinarfunktion, nach der viele Lehrer meinen, daß leistungsunwillige Schüler mit Hilfe schlechter Noten bestraft werden sollen (ZIELINSKI,1973).

Im Gegensatz zu der geringen Beachtung, die die Leistungsbeurteilung in der Lehrerausbildung findet, steht die ihrer Bedeutung angemessene Beachtung durch die einschlägigen Fachwissenschaften. Parallel mit der Ausarbeitung der Fachdidaktiken gehen auch verstärkte Bemühungen um die Lernzielformulierung und um die an test-theoretischen Methoden orientierte lernzielbezogene Leistungsmessung einher. Von dieser Position aus wurden speziell die traditionellen Methoden der Leistungsbeurteilung durch Notengebung in der Literatur zur wissenschaftlichen pädagogisch-psychologischen Diagnostik lange Jahre hindurch scharf kritisiert, etwa von DOHSE, 1963; WEISS, 1965; FLITNER, 1966; GÖLLER, 1968; deGROOT, 1971; INGENKAMP, 1971; ZIELINSKI, 1973; KUTSCHER, 1977, um nur eine kleine Auswahl zu nennen.

Trotzdem hat sich das außerordentlich einfach strukturierte diagnostische Modell, das der Leistungsbeurteilung im schulischen Bereich zugrundeliegt, bisher erfolgreich verändernder Einflüsse aus der umfangreichen Literatur zu dieser Thematik erwehrt.
Übliche Argumente von Seiten der Lehrer gegen die Verwendung einer verbesserten Methodik (etwa der lernzielorientierten Leistungsmessung) verweisen darauf, daß diese Meßverfahren zu unökonomisch seien. Dadurch wird aber ein Nebengütekriterium zu einem Hauptgütekriterium des Meßverfahrens gemacht.

Gegen die Anwendung normgruppenorientierter Leistungsmeßverfahren hingegen wird die geringe Identität und Synchronisierung der Lerninhalte angeführt, vom Niveau der Stoffpläne der Bundesländer bis herunter auf das Niveau der Lehrtätigkeit verschiedener Lehrer des gleichen Faches der gleichen Klassenstufe an ein und derselben Schule.

Ein letztes Argument gegen die Anwendung objektivierter Meßverfahren lautet, daß es im Falle ihrer Anwendung unmöglich sei, "pädagogisch wirkende" Zensuren zu geben. Darunter wird eine weitere subjektive Manipulation an der Rangziffer, mit der die Leistung versehen wird, verstanden - in einem Fall in Richtung auf einen höheren Rang hin, um den betreffenden Schüler zu ermutigen, in einem anderen Fall auf einen niedereren Rang hin, um den Schüler zu verstärkter Anstrengung zu motivieren (etwa im Sinne der "Disziplinarfunktion" KLEBER's,s.o.).Leistungsmessungen haben aber doch in erster Linie wahr zu sein, den tatsächlichen Stand des Individuums im Vergleich zu einer größeren oder kleineren Stichprobe oder zu einer vor der Messung festgelegten Norm widerzuspiegeln, was keineswegs einer zusätzlichen pädagogischen Wertung des im übrigen von dieser Wertung getrennten Ergebnisses der Leistungsmessung eine Absage erteilen soll.

Die Ablehnung, die z.B. lernzielorientierte Leistungstests von seiten der Lehrer erfahren, wird oft doch zu vordergründig ausschließlich mit der Schwerfälligkeit eines konservierenden, statischen Systems, wie es die Schule darstellt, erklärt oder gerade bezüglich der Praxis in den Gymnasien mit dem - vergleichsweise zu den Grund- und Hauptschullehrern - geringen Kontakt der dort tätigen Lehrer mit Pädagogik und Psychologie. Dieses Beharren auf den konventionellen, subjektiven und schwer kontrollierbaren Meßmethoden hat wohl auch noch andere Gründe:

- die Untersuchung der Problematik des Lehrerurteils als Gegenstandsbereich der Pädagogischen Diagnostik, die speziell auch von der Persönlichkeits- und Sozialpsychologie beeinflußt wird, verwendet heute komplizierte statistische Methoden, die sich als Barriere auf die Kommunikation zwischen Erziehungswissenschaftlern und praktisch arbeiten-

den Pädagogen auswirken.

- Analog zum Konstrukt der "Meßfehler"wird in der Forschung meistens zusammenfassend von "Urteilsfehlern" gesprochen, wenn Inkongruenzen zwischen Lehrerurteilen, Lehrerurteilen und Testergebnissen oder systematische Tendenzen in den Urteilen festgestellt werden.
"Fehler" jedoch bedeuten gerade im Bereich der Schule meist selbstverschuldete Abweichungen vom "richtigen" Verhalten, was die Kommunikation zwischen Forschern und Lehrern weiter erschwert. Allerdings ist der Ausdruck "Urteilsfehler", wenn er nicht im statistisch-methodischen Sinne gebraucht wird, hier auch fehl am Platze, denn er würde die Möglichkeit eines richtigen, d.h. höchst validen Urteils voraussetzen - eine Vorstellung, die heute wohl überholt ist.

- Die objektivierte Leistungsmessung an Schülern ist nicht nur auf die Leistung der Schüler bezogen, sondern darüber hinaus auch indirekt Messung der Vermittlungsleistung des Lehrers. Während etwa bei lernzielorientierten Leistungstests der zufallskritische Vergleich der tatsächlichen Schwierigkeitsgrade einzelner items mit den in der Planungstafel zusammengefassten Lernzielen auch eine Aussage über die Güte der Vermittlung der Lerninhalte durch den Lehrer gestattet, wird bei traditionellen Leistungsprüfungen ein unerwartet niedriger Zensuren-Mittelwert der Klasse mit besonders geringer Lernleistung, besonderer Schwierigkeit des Stoffes oder ausnahmsweise der Aufgaben erklärt, keinesfalls aber mit Unzulänglichkeiten des Vermittlungs- oder Beurteilungsverfahrens oder gar der Personen, die dieses durchführen. Diese Tendenz zu externaler Fremdattribuierung bei Versagen der Schüler (aufgezeigt u.a. in Untersuchungen von BECKMANN, 1973; MEYER und BUTZKAMM, 1975; SCHWARZER, 1977) ist bei einem möglichen objektivierten Nachweis geringer Lehreffizienz nicht mehr praktikabel und verliert ihre ich-schützende Funktion für den Lehrer.Eine Bereitschaft zur Verwendung einer solchen Methode ist bei so negativen Nebeneffekten wohl nicht zu erwarten.

- Die Leistungsmessung ist auch Grundlage der gesellschaftlichen Auslese. Die Beurteilung im Schulbereich, von der Grundschule bis zur Universität, ist auch ein Mittel zur Ausübung der sozialen Macht, die von der Gesellschaft an Agenten delegiert wird, die in gesellschaftlichem Auftrag ebendiese Gesellschaft in der neuen Generation reproduzieren sollen durch Verteilung der Sozialchancen über Zensuren und Zeugnisse und den sich daraus ergebenden Prädestinationen für höher- oder geringer qualifizierte Berufe. (ZIEGENSPECK, 1977). Jedem Vorschlag der Veränderung der praktizierten schulischen Beurteilung wird so fast automatisch ein subversives Rütteln an der Gesellschaftsordnung zugeschrieben.

  Gerade dieser letzte Aspekt der gesellschaftlichen Selektion durch Vergabe von Zensuren und Zeugnissen ist es, der die Bedeutung dieses Vorgangs über den rein pädagogischen Raum hinaushebt und ihn zum Gegenstand gesellschaftlichen Interesses werden läßt. Das Bekenntnis zur sozialen Mobilität allein durch Leistung des Individuums, die aus Artikel 12 des Grundgesetzes folgt, ist solange zweifelhaft, wie sachfremde Einflüsse schon zu Beginn der schulischen Sozialisation (CH.SCHWARZER,1976,S.79;s.a.INGENKAMP, 1971; WEISS, 1965) das Schul-Schicksal und damit den weiteren Lebensweg der Kinder vorzuzeichnen beginnen.

Nun haben wir etwas nachvollzogen, was nicht nur im Alltagsleben, sondern auch in der Unterrichtsforschung - soweit sie an den lehrerbildenen Institutionen referiert wird - zu beobachten ist: sobald die Begriffe "Schule" und "Beurteilung" in Verbindung gebracht werden, wird fast ausschließlich die Leistungsbeurteilung diskutiert.

Im Schulbereich wird aber nicht nur die Leistung der Schüler beurteilt:

1. Sieht man sich die Zeugnisformulare an, so bemerkt man auch wenigstens zwei Beurteilungskategorien, die auf eine umfassende Beurteilung nicht-fachspezifischer Inhalte hindeuten und die etwa als "Betragen", "Verhalten" und "Fleiß",

"Mitarbeit" bezeichnet werden. Erfahrungsgemäß weisen diese Kategorien in der Praxis nur drei oder vier Ausprägungsgrade auf im Gegensatz zu der in Zeugnissen üblichen Skalierung mit 6 Notenschnitten oder gar zu der bei Zensuren währen des Schuljahres angewandten Skalierung mit 11 bzw. 21 Notenschritten.

2. Seit kurzer Zeit ist die Schulbehörde in einigen Bundesländern wenigstens für die beiden ersten Grundschulklassen von der Beurteilung anhand von Ziffern abgegangen und verlangt jetzt von ihren Lehrern eine Art von Persönlichkeitsgutachten mit besonderer Berücksichtigung des Leistungsaspektes. Inzwischen kann man aber den Äußerungen von Eltern bei Schulpsychologischen Beratungsstellen entnehmen, daß ein Teil der Lehrer diese Gutachten so mit pädagogischen und psychologischen Termini füllt, als sei der Adressat ein Schulpsychologe, womit diese Gutachten für ein Großteil der Eltern einfach unverständlich werden. Eine weitere Gruppe von Lehrern hält sich im Aufbau und Inhalt der Gutachten sehr eng an die soeben abgeschafften Zeugnisformulare und verwendet nur statt der Ziffern 1 - 6 die verbalen Bezeichnungen "sehr gut" bis "ungenügend". Diese Erscheinungen mögen einmal darauf basieren, daß den in der Erstellung von Persönlichkeitsgutachten nicht besonders bewanderten Lehrern von der Schulbehörde kaum Hilfsmittel in die Hand gegeben werden, zum anderen darauf, daß die Lehrer gewisse Bedenken haben, mit expliziten Persönlichkeitsgutachten an die Öffentlichkeit zu treten.
All das ist einsehbar, wenn man bedenkt, wie schwierig es für die nicht ausgebildeten Lehrer ist, ein Persönlichkeitsgutachten über 6-8-jährige Kinder zu formulieren, das über die bloße Beschreibung besonders häufig beobachteter Verhaltensweisen hinausgehen soll. Außerdem wird durch die Gutachtenerstellung die grundsätzliche Problematik von Urteilen keineswegs aufgehoben - es wird nur die Problematik der Leistungsbeurteilung durch die der Persönlichkeitsbeurteilung ersetzt.

3. In einzelnen Entscheidungssituationen über die weitere Schullaufbahn der Schüler werden ebenfalls über die reine Leistungsdiagnostik hinausgehende Urteile erwartet: bei der Überweisung an die Sonderschule, Einschulung, Empfehlung für eine weiterführende Schulart und der Entscheidung über das Verbleiben im Gymnasium nach dem Probehalbjahr.

4. Bis vor einiger Zeit forderte die Schulbehörde von Lehrern in einzelnen Bundesländern die permanente Persönlichkeitsbeurteilung von Schülern anhand von "Schülerbogen". Das Anlegen von "Schülerbogen" wurde zur Auflage gemacht, zu deren Führung nicht weniger als planmäßige Beobachtung der Schüler und, darauf aufbauend, eine umfassende Beschreibung der Schülerpersönlichkeit gefordert wurde (s. auch ULICH/MERTENS,1973,S.11 f). Die Hilfestellungen, die von der pädagogischen Diagnostik geleistet wurden (ENGELMAYER, 1960; FEIGEL u. KEITEL, 1964; DONAT, 1965; KAIPF, 1968), zeichnen sich durch einen Mangel an theoretischer Fundierung, gängige populärpsychologische Begründungen der verwendeten Kategoriensysteme und gehäuftes Vorkommen populärer Stereotypen aus.
In der genannten Literatur treffen wir auf naiv-psychologische Theorien über die Persönlichkeit von Schülern. Im folgenden Kapitel werden wir uns mit ebensolchen naiv-psychologischen Wissensbeständen der angesprochenen Zielgruppe "Lehrer" beschäftigen, die allerdings nicht so explizit formuliert vorliegen wie in der angesprochenen Literatur dieses Abschnitts.

2. Naive Persönlichkeitstheorien von Lehrern

Der Alltagsmensch besitzt in seinem naiv-psychologischen Wissensbestand eine Anzahl von Konzepten bezüglich seiner Mitmenschen, die sowohl deren Verhalten als auch deren Persönlichkeit als Verhaltensquelle betreffen .(LAUCKEN, 1974, S.30).
LAUCKEN (1974) identifiziert diese Wissensbestände, die er vorwiegend unter dem Verhaltensaspekt analysiert (naive Verhaltenstheorie), als das Wissen, das man "im tagtäglichen Umgang mit Mitmenschen benutzt, um deren Verhalten zu beschreiben, zu erklären und zu antizipieren" (a.a.O., S.18), und kategorisiert diese naiven Erklärungskonzepte in solche,

- die bewußt und gewollt hervorgebracht erscheinendes Verhalten erklären sollen und solche, die
- "das Verhalten einer Person gleichsam als 'automatische' Reaktion auf bestimmte, der Kontrolle der betreffenden Person entzogene Faktoren" deutend erklären sollen. (LAUCKEN, 1974, S.59).

Ferner gruppiert der genannte Autor das Gesamt der naiv-psychologischen Konzepte in

- "Alltagstheoretische Konstruktionen, welche ein angenommenes aktuell ablaufendes Geschehen erfassen" (Prozeßkonstrukte), und
- "Alltagstheoretische Konstruktionen, die einer Person attribuierte überdauernde Bestände meinen". (Dispositionskonstrukte). (LAUCKEN, 1974, S.61).

Gemäß unserem Vorhaben, die Beurteilung der Schülerpersönlichkeit durch Lehrer zu untersuchen, werden wir uns bei den folgenden Überlegungen nicht auf so hohem Abstraktionsniveau bewegen können wie LAUCKEN (1974), sondern auf einem "teilkonkretisierten" Niveau (ebd., S.36), auf dem sich subjektbezogene Individualtheorien bezüglich klassifizierter Untersuchungsobjekte befinden, d.h. es sollen Theorien einer Personengruppe bezüglich einer anderen Personengruppe untersucht werden.

Eine weitere Einschränkung muß erfolgen durch das Ausklammern der Prozeßkonstrukte und das Konzentrieren auf Dispositionskonstrukte.

Der eingangs angeführte Satz muß nach diesen Einschränkungen modifiziert also lauten: Lehrer besitzen in ihrem naiv-psychologischen Wissensbestand eine Anzahl von Konstrukten bezüglich der Persönlichkeit von Schülern.
Sucht man nach konkreter formulierten Inhalten dieses Wissens, so stößt man auf Annahmen

"1. über 'die Natur des Menschen', über Charaktertypen usw"., z.B.: "...Hätten's gar nicht nötig, daß die Mutter arbeitet (!), aber ein Wagen war nötig. Die kommen an wie eine gräfliche Familie...Soweit ich informiert bin, soll sie mittags oft Jungen in der Wohnung haben...Die Eltern leugnen das zwar, aber das paßt zu deren Charakter". (HÖHN, 1967, S.79).

2. über das Vorkommen von bestimmten Eigenschaften bei bestimmten Leuten-vom nationalen Vorurteil bis zum Stereotyp des 'schlechten Schülers': z.B.: "schlechte Schüler sind unzuverlässig und verschlagen". (HÖHN, 1967, S.92).

3. über bestimmte Ursache-Wirkungsverhältnisse zwischen Verhaltensäußerungen und dem zugrundeliegenden "Charakter": z.B.: "Charakterlich schlecht; er ist sehr faul und lügenhaft; versucht damit seine Faulheit zu decken". (HÖHN, 1967, S.60).

4. über ebenfalls 'gesetzmäßige' Zusammenhänge zwischen bestimmten Aktionen und Reaktionen (bei Faulen hilft nur Druck, usw.) (ULICH/MERTENS, 1973, S.102).

Das Urteilen, das auf diesem Wissen basiert, folgt aus der sozialen Interaktion Lehrer-Schüler und wirkt wieder in diese hinein (KLEBER, 1978, S.599). Allerdings existiert noch keine umfassende Theorie sozialen Handelns, die es gestatten würde, das in den Rahmen sozialer Interaktion eingebettete Urteilen zu erklären. Wir müssen uns daher bei der Analyse von Urteilsprozessen zunächst einmal themenverwandten wissenschaftlichen Konstrukten zuwenden, deren

Begriffs-Vokabular wir bei dieser Analyse zu verwenden gedenken. Darunter fallen die Problemkreise des Eigenschaftsbegriffes, der sozialen Wahrnehmung und daraus folgend die Problemkreise der Erwartungseffekte, der Kausalattribution und der Selektivität.
Zum Schluß soll noch auf die "impliziten Persönlichkeitstheorien" als den eigentlichen Gegenstand der Untersuchung eingegangen werden.

II Problemkreise mit Bezug zur 'impliziten Persönlichkeitstheorie'

2.1. Das Problem der Eigenschaften

Es wäre doch recht umständlich und nicht sehr ökonomisch, in der Kommunikation zur Kennzeichnung eines Dritten alle Situationen und alle Verhaltensweisen dieses Dritten in diesen Situationen zu beschreiben. Einfacher und ökonomischer ist es, tendenziell gleichgerichtet erscheinende Verhaltensweisen zu bündeln und mit einem Etikett zu versehen, das stellvertretend für diese Verhaltensweisen steht.
Daß sich die Kommunikationspartner mit Hilfe dieser Etiketten verständigen können, obwohl zwischen ihnen eventuell unterschiedliche individuelle Auffassungen darüber bestehen, ob Häufigkeit, Intensität und Konsistenz des beobachteten Verhaltens dieses Dritten eine Kennzeichnung mit gerade diesem Etikett rechtfertigen, erklärt LAUCKEN (1974, S.24 f) mit der Existenz zwar unterschiedlicher Individualtheorien, die jedoch Teile übergreifender Rahmentheorien sind,
Der Alltagsmensch weiß daher um die Individualtheorie des anderen, ohne daß die Kommunikationspartner sich diese gegenseitig erklärt haben. LAUCKEN (1974,S.25) gelangt zu der Ansicht, "daß das alltagsgebräuchliche Repertoire an Erklärungskonzepten, seien diese nun rahmentheoretischer oder spezifischer Art, aus Begriffen der Alltagssprache besteht, so daß eine der betreffenden Sprache mächtige Person auch um jene Konzepte weiß, die außerhalb der Individualtheorie liegen" - eine Argumentation, die sich eng an WHORF anschließt. (siehe dazu auch die Diskussion von HÖRMANN, 1970, S.334-351).

Als Leistungen dieser "Be-Eigenschaftung" führen ULICH/MERTENS (1973, S.27 f) auf:

- "Die Verwendung von Eigenschaftsbegriffen erleichtert uns die Verständigung über unsere Mitmenschen".
- "Die 'Kenntnis' einiger weniger Eigenschaften einer anderen Person erlaubt schon eine ganzheitliche Persönlichkeitsbeurteilung".
- "Die Kenntnis von Eigenschaften schafft in uns Erwartungen im Hinblick auf das (künftige) Verhalten der Person in gleichartigen Situationen."

Mit der Verwendung von Eigenschaftsbezeichnungen verbunden sind jedoch Implikationen, die die mittels Eigenschaften gefällten Urteile verzerren, d.h. die Validität der Urteile sehr gering halten: (ULICH/MERTENS,1973,S.34 ff)

- Annahme einer Konsistenz und Konstanz des Verhaltens.
  Bei der Verwendung von Eigenschaftsbezeichnungen im Sinne konstanter Verhaltensbereitschaften entledigt sich der Urteiler leicht der Notwendigkeit, andere z.B. situative Einflußfaktoren zu berücksichtigen. Diese Annahme, alltagssprachlich formuliert: Ein Mensch verhält sich in vergleichbaren Situationen gleich, könnte sich in folgender konkreten Lehrererwartung ausdrücken: Der Schüler hat in der letzten Klassenarbeit abgeschrieben und wird das in der heutigen Klassenarbeit auch wieder tun.
- Annahme der Generalität von Eigenschaften.
  Der Benutzer des Eigenschaftsbegriffs nimmt an, daß sich eine Person in verschiedenen Situationen gleich verhält: Der Schüler A war ehrlich, als er gestern sofort zugab, daß er diesen Streich gespielt hat; also wird er auch heute ehrlich sein und die gefundenen 100,- DM beim Fundbüro abgeben.
  Diese Annahme der Unabhängigkeit des Verhaltens von unterschiedlichen Situationen und Rollen trifft in besonderem Maße auf Verhaltensweisen zu, die mit der Anerkennung von Wertsystemen verbunden sind (Moral, Ideale usw). und verleitet den Urteiler, die eventuelle situationsbedingte Inkonsistenz der Anerkennung dieser Wertsysteme nicht zu berücksichtigen.

- Annahme der Universalität von Eigenschaften.
  Grundsätzlich kann man nach dieser Annahme alle Menschen hinsichtlich aller möglichen Verhaltensdispositionen beurteilen. Dabei ist jedoch die soziokulturelle Relativität von Urteilskategorien, Wertsystemen und Erziehungsidealen zu bedenken. Nicht nur der Urteilende definiert die Situation, in der er den Handelnden vorfindet und dessen Verhalten er auf eine Eigenschaft rückbezieht, sondern auch der Handelnde definiert diese Situation - eventuell aber ganz anders als der Urteiler.

  Ein Beispiel dazu: Ein dreijähriges Kind erzählt eine Wunschvorstellung als real erlebte Geschichte. Der erwachsene Zuhörer kann nun von seiner Vorstellung von Ehrlichkeit her dem Kind die Eigenschaft "lügenhaft" zuschreiben; die Wertvorstellung,aus der heraus diese Be-Eigenschaftung erfolgt, hat jedoch für das Kind selbst, dem sich Wunsch und Wirklichkeit unidentifizierbar vermischen, keine Gültigkeit.

Eigenschaften sind also keineswegs universal zu verwenden; anknüpfend an die unterschiedlichen Situationsdefinitionen sind sie vielmehr "Stellungnahmen in den erlebten Beziehungen zwischen sich und der sozialen Umwelt...Sowohl die (Eigenschafts- d.Verf.) Bezeichnungen selbst als auch ihre Verwendungsweise sind Produkte sozialer Wechselwirkungen, also Ergebnisse von Lernprozessen." (ULICH/MERTENS, 1973, S.35).

Das Bedürfnis des Alltags-Diagnostikers, in seinen Urteilen auf 'Eigenschaften' als überdauernde Verhaltensdispositionen abzuzielen, ergibt sich daraus, daß Personen nicht als passive Objekte erlebt werden, sondern als agierende und reagierende Handlungszentren, die tatsächliche oder mögliche Interaktions- und Kommunikationspartner sind. Bleibt die Diagnostik auf dem Niveau bloßer Beschreibung der hochvarianten Verhaltensweisen anderer Personen, so kann der Diagnostiker keine sinnvollen eigenen Handlungspläne entwerfen, er wird vom Verhalten des Gegenübers gewissermaßen überrascht. Es ist für ihn daher notwendig, sich einen sinnvollen Handlungsraum zu entwerfen, der strukturiert, wenigstens für eine bestimmte Zeit stabil und bedeutsam für ihn

ist. Er muß daher aus vergangenen und gegenwärtigen Beobachtungen von Ereignissen der sozialen Umwelt zukünftige Ereignisse antizipieren und das mit einem gewissen Anspruch auf tatsächliches Eintreffen dieser Ereignisse. Optimal geschieht das, indem die unübersehbare Anzahl von Informationen über äußere Reize und Reizkonstellationen (Verhaltensweisen sowie Verhaltensprodukte) reduziert wird, indem ihm in der Tendenz gleichgerichtet erscheinende zusammengefaßt und auf eine 'in' der Person vorhandene, latente Verhaltensdisposition (Eigenschaften) zurückgeführt werden. (siehe dazu: LUKASCZYK, 1969, zit.in.ULICH/MERTENS,1973, S.67; HASTORF,1970).
Für den Vorgang der Attribuierung solcher Eigenschaften bei aktuellem Urteilen bietet BRUNSWIK (1956) ein Modell an.

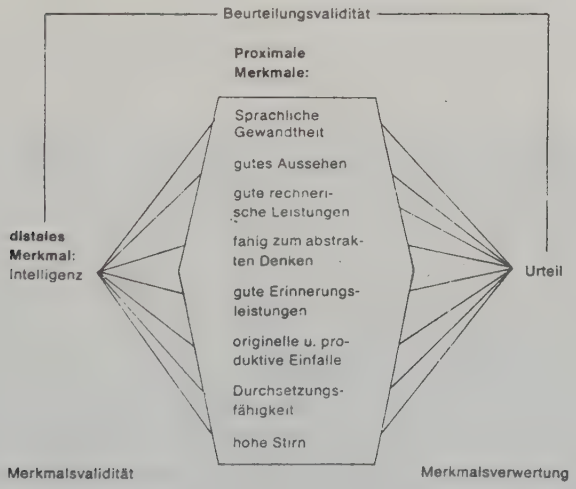


Abb.1: Linsenmodell nach BRUNSWIK (1956); aus: ULICH/MERTENS (1973, S.75).

In Bezug auf das Urteilen in der Lehrer-Schüler-Beziehung ist aus diesem Modell folgendes zu entnehmen:
Wenn ein Lehrer zu einem Urteil über Persönlichkeitsmerkmale eines Schülers kommen will, das er entweder für interne (Handlungsentwurf) oder externe Zwecke (Besprechnung mit anderen Lehrern oder mit Eltern, Erstellung eines Gutachtens) benötigt, so kann er das Merkmal selbst nicht beobachten. Er sucht also in Verhalten, Verhaltensprodukten, physischen und sozio-ökonomischen Merkmalen des Schülers beobachtbare Indikatoren ("proximale Merkmale"), die ihm gemäß seines naiv-psychologischen Wissens eine Aussage (Urteil) über den wahrscheinlichen Ausprägungsgrad des zu erfassenden Konstrukts ("distales Merkmal") erlauben.

Statistisch gesehen, gestattet uns die in Lehrerurteilen auftretende Korrelation zwischen einem beobachtbaren proximalen Merkmal und dem Urteil über das zu erfassende Konstrukt eine Aussage über das Ausmaß der "Merkmalsverwertung".

> Korrelieren in Lehrerurteilen die Ratings der Kategorien 'hohe Stirn' und 'Intelligenz' zu r = +.98, so kann man sagen, daß die Lehrer bei der Beurteilung der Intelligenz ihrer Schüler in praktisch jedem Fall die Stirnhöhe der Schüler als einen Indikator verwenden.

Die Frage ist hier nur, ob dieser Indikator auch valide für die Beurteilung dieses Konstrukts ist (Frage nach der "Merkmalsvalidität"). Die Merkmalsvalidität kann festgestellt werden, indem proximales Merkmal und Konstrukt objektiviert gemessen werden - sofern möglich. Die Korrelationen beider Messungen in einer hinreichend großen Stichprobe geben an, wie genau das Urteil ist, d.h. wie präzise das Urteil die Realität abbildet. Je invalider die Indikatoren für das zu beurteilende Konstrukt sind, desto weniger zutreffend wird das Urteil über dieses Konstrukt sein und umgekehrt.

> Das obige Beispiel fortführend, könnten Messungen der Stirnhöhe und die Ergebnisse parallel dazu durchgeführter Intelligenz-Tests zu r = .oo korrelieren; ein Vergleich mit dem Koeffizienten der Merkmalsverwertung zeigt, daß dieser Indikator, den die Lehrer mit hoher Übereinstimmung bei der Intelligenzdiagnose verwenden, dennoch invalide für diese Diagnose ist.

Das aufgeführte Beispiel zur Feststellung der Merkmalsvalidität birgt jedoch ein Legitimationsproblem in sich: während die Messung der meisten proximalen Merkmale durchaus noch so objektiviert erfolgen kann, daß sie ein realitätsgetreues Abbild der Tatsachen darstellen mag, wird im Fall der Messung des Konstrukts ja nur das vom Lehrer definierte Konstrukt durch ein wissenschaftliches Konstrukt ersetzt, dessen Wert als valides Außenkriterium auch nur innerhalb der Theorie unbestritten ist, die dieses Konstrukt hervorgebracht hat. Fragebogen und Tests bringen zwar einen Gewinn an Objektivität, sind jedoch mit Validitäts- und Reliabilitätsproblemen beladen, die man gerade vermeiden wollte; Expertenurteile als Außenkriterium bringen wiederum Objektivitätsmängel in den Meßvorgang ein.

Diese Validitäts-Problematik, nach der ein Urteil schon allein dann invalide werden muß, wenn darin Eigenschaftsbegriffe verwendet werden, führten speziell bei Kritikern der Persönlichkeitsbeurteilung von Schülern durch Lehrer zu der Forderung, bei Beurteilungen im Schulalltag und in Untersuchungen zu dieser Thematik keine Eigenschaftsbezeichnungen mehr zu verwenden, sondern Situations-Verhaltensitems (KLEITER, 1973; SCHWARZER/SCHWARZER,1977), die eine urteilsfreie Beschreibung des Verhaltens ermöglichen sollen. Auch dieser Vorschlag ist nicht ganz unproblematisch, wie die Diskussion um das Bandbreite-Fidelitäts-Dilemma, die Inferenzproblematik und die Legitimationsprobleme bei der Auswahl der 'richtigen' Verhaltensweisen für die zu beurteilenden Verhaltensdispositionen zeigen. (s.a. SCHWARZER/SCHWARZER, 1977, S.42-79).
Die Schwierigkeiten bei der Bestimmung der Urteils-Validität führten zu einer Verlagerung der Forschungs-Interessen auf Probleme der Merkmalsverwertung (z.B. SCHWARZER, 1976), der Urteilstendenzen (z.B. HOFER, 1974) oder des Auffindens von Beziehungen zwischen Persönlichkeitsmerkmalen, Verhalten und Urteil (z.B. FIEDLER, 1964).

2. Die soziale Wahrnehmung

Alle Informationen über unsere Umwelt - und hier soll speziell auf die soziale, nicht die dingliche Umwelt eingegangen werden - stammen ursprünglich aus Beobachtungen. Diese zunächst noch objektiven Daten, die physikalischer Natur sind und durch das Einwirken von Reizen auf das Sensorium entstehen, treffen allerdings nicht auf neutrales Gebiet, sondern auf ein Wesen, das schon eine soziale Lerngeschichte hat; es hat schon Erfahrungen im Zusammenhang mit solchen Daten, die sich in Erwartungen gegenüber diesen Informationen niederschlagen und die das Individuum die Informationen nach Maßgabe diese Erfahrung bewerten lassen. (BRUNER, 1951). Es ist dem Individuum also unmöglich, ein realitätsgetreues Abbild der Wirklichkeit zu bekommen; es erhält vielmehr ein individuell wahrgenommenes. Das Wahrnehmen zeigt sich hier nicht als passivrezeptiver, sondern als aktiv gestaltender Vorgang, die Wahrnehmung selbst als Kompromiss zwischen der objektiven Information und den Erwartungen des Beobachtenden (BRUNER, 1951).
Dem Wahrnehmenden ist es nicht möglich, objektive Informationen und wertende Stellungnahme zu trennen, selbst wenn er versucht, sich den Wahrnehmungsvorgang bewußt zu machen, da sich seine Aufmerksamkeit den Umweltobjekten nicht in ihrer physikalischen Form zuwendet, sondern in ihrer erlebten Bedeutung. (s. ULICH/MERTENS, 1973, S.61).
WARR und KNAPPER (1968) haben versucht, in einem Modell des Vorgangs der Personwahrnehmung die beteiligten Variablen zueinander in Beziehung zu setzen. (s.Abb.2).

Der Wahrnehmungsprozeß läuft nach diesem Modell folgendermaßen ab:
Die Informationen, die zum Urteil beitragen, sind aktuelle Informationen über die Reizperson, gespeicherte Informationen über die Reizperson (Erfahrungswissen, Erwartungen, Stereotype, etc.) und aktuelle Kontext-Informationen. Diese Informationsquellen stehen miteinander in Wechselwirkung, beeinflussen einander.

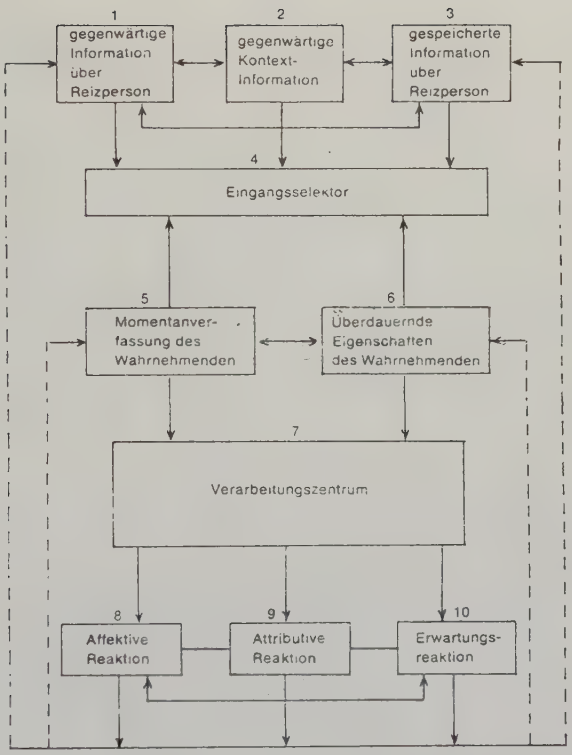


Abb. 2: Modell der sozialen Wahrnehmung nach WARR u. KNAPPER, 1968, S.2o; übers.n. GRAUMANN, 1969,S. 39.

(aus: ULICH/MERTENS,1973, S.7o)

Gemäß der momentanen Verfassung des Wahrnehmenden (z.B. emotional-affektive Zustände) und seiner Reaktionstendenzen (um die soeben geführte Diskussion über Eigenschaftsbegriffe sogleich nutzbringend zu verwenden, soll hier die Übersetzung GRAUMANNs (1969, S.39) als "überdauernde Eigenschaften" nicht verwendet werden; gemeint sind wohl eher Haltungen, Einstellungen,etc.)werden die eingehenden Informationen selegiert und dem Verarbeitungszentrum zugeleitet. Im Verarbeitungszentrum sind vor allem zwei Regeln der Verarbeitung zu finden:

- Inferenz-Regeln, die Analogieschlüsse von einem beobachteten Merkmal auf ein anderes, nicht beobachtetes Merkmal ermöglichen (siehe auch Kapitel 'Implizite Persönlichkeitstheorien') und
- Kombinationsregeln, die die Informationen nach bestimmten Strategien zusammenfassen. (s. auch ULICH/MERTENS, S. 1973, S.69).

Das Produkt der Verarbeitung ist das Urteil in seinen sich gegenseitig beeinflussenden Komponenten

- affektive Reaktion:(emotional-wertende Stellungnahme gegenüber der Reizperson)
- attributive Reaktion:(Zuschreibung von Verhaltensdispositionen, Eigenschaften aufgrund von Schlußfolgerungen aus dem Verhalten der Reizperson)
- Erwartungsreaktion: (Erwartung bezüglich zukünftigen Verhaltens der Reizperson).

Das den Wahrnehmungsprozeß vorläufig beschließende Urteil wirkt nun seinerseits wieder zurück auf einzelne Faktoren des Wahrnehmungsprozesses - es beeinflusst die Faktoren der Informationsebene, die Momentanverfassung und die überdauernden Persönlichkeitsmerkmale des Wahrnehmenden. Zu weitgehend scheint die Interpretation dieser Rückkoppelungen von ULICH/MERTENS, die dazu bemerken, es sei "ja einleuchtend, daß das Urteil das konkrete Verhalten der beurteilten Person selbst verändern kann" (a.a.O., S. 71). Gerade das leistet das Modell von WARR und KNAPPER (1968) nicht, nämlich eine Einbettung des Wahrnehmungsprozesses und des diesem Prozeß vorerst beschließenden Urteils in einen interaktiven Zusammenhang, in einen Ablauf aufeinanderfolgender Aktions-Reaktions-Sequenzen, wie ihn z.B. KAMINSKI (1959,S.110,111) beschreibt.

2.3. Self-fulfilling Prophecies

Wie wir gesehen haben, ist das Modell des Wahrnehmungsprozesses von WARR und KNAPPER (1968) in dem Sinne a-sozial, als es die Auswirkungen des Urteils auf die soziale Umwelt, speziell auf den Beurteilten nicht behandelt.
Zur Weiterführung dieser Thematik knüpfen wir besonders an die Erwartungsreaktion als einer Komponente des Urteils an. Diese Erwartungsreaktion ist prognostischer Natur: urteilsgemäß wird ein ganz bestimmtes zukünftiges Verhalten vom Beurteilten erwartet. Diese Erwartung ist nun wiederum bei auf das Urteil folgenden Kontakten des Urteilers mit dem Beurteilten handlungsrelevant für das Verhalten des Urteilers. Das Urteilerverhalten wirkt wiederum beim Beurteilten als Information über eine Reizperson, die in diesem Fall der Urteiler darstellt. Der weiter oben dargestellte Wahrnehmungsprozeß läuft wieder ab - diesmal auf der Seite des Beurteilten, der jetzt zum Urteiler wird.Man könnte sich die wechselseitig aufeinander bezogenen Wahrnehmungsprozesse als jeweils um eine Phase verschoben zu beiden Seiten einer Symmetrie-Achse aus Verhalten und Verhaltensprodukten der beiden Interaktionspartner gruppierte Sequenz vorstellen.

Die Auswirkungen eines speziellen Erwartungseffektes, der 'Self-fulfilling prophecies', (MERTON, 1948), hat besonders ROSENTHAL (1963,1964,1965,1966,1967,1970); ROSENTHAL u. ANDERSON (1969); ROSENTHAL u. EVANS (1968); ROSENTHAL u. FODE (1963,b); ROSENTHAL,FRIEDMAN u. KURLAND, (1966); ROSENTHAL u. JACOBSON (1966,1968); ROSENTHAL u. LAWSON,(1964) zunächst mit Tieren, dann mit Versuchspersonen im Labor als Reizobjekten und zuletzt in der Realsituation der Schule untersucht.
Im Grundmodell aller Untersuchungen von ROSENTHAL werden im Urteiler falsche Erwartungen bezüglich der Reizperson(en) induziert, die nach einer gewissen Interaktionszeit von Urteilern und Reizpersonen dazu führten, daß das Erwartete tatsächlich eintrat.

Aus den Untersuchungen unter Laborbedingungen folgerte ROSENTHAL (1967), daß diese Erwartungen in kleinsten Hinweisreizen (covert communication) des Erwartenden zum Ausdruck kommen, da der 'Pygmalion-Effekt' (ROSENTHAL,1968) zwar auch dann auftrat, wenn dem Erwartenden das Verhalten genau vorgeschrieben war und kontrolliert wurde, jedoch nicht auftrat, wenn die direkte Kommunikation zwischem dem Erwartenden und dem Erwartungsobjekt unterbunden wurde.

Bei den Experimenten im Schulbereich war jedoch besonders die Interaktion zwischen den erwartenden Lehrern und den Schülern so wenig kontrolliert, daß forschungsmethodische Kritik angebracht erscheint (s.ELASHOFF u. SNOW,1972). Inzwischen stützen jedoch weitere Untersuchungen (z.B. BROPHY u. GOOD, 1976)die Annahme solcher Erwartungseffekte.
Die Ansicht,der 'Pygmalion-Effekt' sei ein stets funktionierender Mechanismus, kann allerdings nicht aufrecht erhalten werden. (HECKHAUSEN,1974; HANKE u. MANDL, 1975).

Diese Darstellung als Mechanismus finden wir auch bei SCHWARZER u. SCHWARZER (1977, S.37), die den Prozeß der 'self-fulfilling prophecy' im Rahmen der Schule erklären:
(Stufe 4): "Die Schüler übernehmen die Erwartungsinhalte in ihr Selbstbild. Aufgrund ihrer Wahrnehmung der Lehrererwartungen korrigieren Schüler die impliziten Hypothesen, die zuvor für ihre Selbstwahrnehmung verbindlich waren. Sie entwickeln ein Selbstkonzept, das sich im kognitiven Gleichgewicht mit dem Fremdbild befindet, über das der Lehrer von ihnen verfügt."
(Stufe 5): "Die Schüler verhalten sich so, wie der Lehrer es von ihnen erwartet. Die Summe der an sie herangetragenen Erwartungen macht die Rolle der Schüler aus. Die meisten Schüler verhalten sich rollenadäquat. Erwartungen, Selbstbild und Verhalten stimmen überein."
(Stufe 6):" Der Lehrer nimmt das erwartungsgemäße Verhalten wahr. Der Kreis hat sich geschlossen: der Lehrer kann seine Erwartung als bestätigt ansehen, denn die Schüler widersprechen seiner Prophezeiung nicht. Je öfter dieser Prozeß stattfindet, um so 'berechtigter' wird die Erwartung und um so unwahrscheinlicher wird ein erwartungswidriges Schüler-

verhalten."

Bedenken können bezüglich des "Funktionierens" der Stufe 5 des geschilderten Ablaufs angemeldet werden. Zunächst einmal unterscheiden die Autoren nicht zwischen der generell in dieser Gesellschaft und in dieser Zeit gerade in der Schule gültigen Lehrer-Erwartungen bezüglich Schülern, (Rolle), und individuellen Erwartungen eines Lehrers bezüglich eines Schülers (Erwartungsreaktion als Komponente des Urteils über diesen speziellen Schüler).
Während sich die gesellschaftlich eingebundenen Rollenerwartungen sicher nur sehr langfristig verändern, können sich die Erwartungsreaktionen eines Lehrers gegenüber einem Schüler (die zusammen mit den attributiven Reaktionen das 'Bild des Schülers' darstellen) relativ zu den Rollenerwartungen wesentlich kurzfristiger verändern.

> Nehmen wir den Fall, in dem ein Lehrer aufgrund gespeicherter Information durch einen Dritten überhöhte Erwartungen einem Schüler mit tatsächlichen Fähigkeits- oder Wissensmängeln gegenüber hegt. Das tatsächliche Verhalten dieses Schülers muß dann beim Lehrer das Erleben einer kognitiven Dissonanz (FESTINGER, 1963) auslösen. Zunächst wird die auf den Eingangsselektor wirkende gespeicherte Information eine Unterdrückung oder Verringerung der Dissonanz bewirken. Je größer aber die Dissonanz ist und je häufiger sie in den aktuellen Informationen über die Reizperson vom Lehrer erlebt wird, desto größer wird der Druck, diese Dissonanz zu beheben, wenn er nicht die Verbindung zur Wirklichkeit verlieren will. Die Urteilsreaktionen müssen sich also den aktuellen Informationen annähern; über die Rückkoppelung des Urteils zur gespeicherten Information wird diese letztere soweit modifiziert, bis keine Diskrepanz mehr zwischen aktuellen und gespeicherten Informationen erlebt wird, was allerdings nicht besagt, daß das Urteil jetzt den objektiven Tatsachen entspricht. Ein gleicher Ablauf ist anzunehmen, wenn sich ein Lehrer im Verlauf häufiger Kontakte ein Urteil über einen Schüler gebildet hat und dieser Schüler nun aufgrund von Fremdeinwirkung sein Verhalten mehr oder weniger plötzlich verändert (familiäre Probleme; Druck der Eltern auf den Schüler, sein Schulverhalten zu ändern; Verhaltensänderung aufgrund der Gefahr, sitzenzubleiben).

In all diesen Fällen von Verhaltensänderungen der Schüler, die im Schulalltag recht häufig sind, wird sich die Wahrnehmung des veränderten Verhaltens durch den Lehrer erst allmählich 'akkomodieren'.
Diese Beispiele zeigen, daß der oben bereits zitierte Satz von SCHWARZER u. SCHWARZER (1977, S.37): " Die Schüler verhalten sich so, wie der Lehrer es von ihnen erwartet" keine durchgängige Gültigkeit hat, speziell auch nicht mit der darauffolgenden rollentheoretischen Begründung der Autoren. Eine einleuchtendere Begründung für dieses Phänomen wäre vielleicht, daß der Lehrer einem Schüler Verhaltensfreiräume je nach seinen Erwartungen gewährt.

> Meldet sich ein als 'gut' beurteilter Schüler auf eine Frage und gerät bei der Antwort ins Stocken, so wird der Lehrer eher die Tendenz haben, abzuwarten, bis der Schüler seine Gedanken gesammelt hat, weil er ja erwartet, daß der Schüler die Antwort weiß. Schließlich gibt der 'gute' Schüler die richtige Antwort.
> Meldet sich hingegen ein als 'schlecht' beurteilter Schüler, der die Antwort ebenfalls weiß und gerät ins Stocken, wird der Lehrer eher die Tendenz haben, abzuwinken, da er ja erwartet, daß er die Antwort nicht weiß, daß nichts Verwertbares mehr kommt. Der 'schlechte' Schüler kommt nicht dazu, seine Antwort zu äußern.
> In beiden Fällen ergibt sich eine Bestätigung der Lehrererwartung, die der Lehrer durch sein Verhalten selbst herbeigeführt hat.

Neuere Untersuchungen legen die Vermutung nahe, daß es individuelle Unterschiede zwischen Lehrern gibt in der Vermittlung ihrer Erwartungen.
BROPHY u. GOOD (1976, S.160 ff) kategorisierten Lehrer nach ihrem Verhalten gegenüber individuellen Differenzen bei Schülern in 'proaktive, reaktive und überreaktive' Lehrer.

Proaktive Lehrer nehmen Unterschiede differenziert wahr, womit die Möglichkeit einer flexiblen Wahrnehmung von Verhaltensänderungen der Schüler gegeben ist und somit auch sich aktuellen Informationen flexibler anpassende Erwartungsreaktionen. Die Wahrscheinlichkeit des Auftretens von 'self-fulfilling prophecies' ist bei proaktiven Lehrern geringer als bei Vertretern der beiden anderen Kategorien.

Der reaktive Lehrer verhält sich gegenüber individuellen Unterschieden wie gegenüber schicksalshaft festgeschriebenen Realitäten, obwohl er diese individuellen Differenzen durchaus wahrnimmt. Dadurch fixiert er auch seine Erwartungsreaktionen stärker als der proaktive Lehrer, weswegen die Wahrscheinlichkeit des Auftretens von Erwartungseffekten bei ihm auch größer ist.
Der überreaktive Lehrer nimmt individuelle Unterschiede der Schüler nicht differenziert wahr, da er frühzeitig in der Interaktion Eigenschaften attribuiert und stereotypisiert. Er fällt vorrangig dichotome Urteile (ordentlich oder unordentlich; faul oder fleißig). Analog dazu sind auch seine Erwartungen rigide, die Wahrscheinlichkeit des Auftretens von 'self-fulfilling prophecies' ist bei ihm am höchsten.

In einer abschließenden pädagogischen Bewertung der Auswirkungen von 'self-fulfilling prophecies' im Schulalltag ist die fördernde Wirkung positiver Lehrererwartungen hervorzuheben; bei negativen Erwartungen hingegen besteht die Gefahr einer Stigmatisierung von Schülern, die besonders in den ersten Jahren der Sekundärsozialisation langfristig negative Auswirkungen auf Einstellungen der Betroffenen bezüglich Leistung, Lernen und Schule zur Folge haben kann.
Es könnte also durchaus als eine prophylaktische Strategie in der Lehrerausbildung angestrebt werden, in Studierenden grundsätzlich positive - keinesfalls aber überzogene - Erwartungshaltungen bezüglich ihrer späteren Schützlinge zu induzieren.

2.4 Selektivität der Wahrnehmung

In jedem der vorhergehenden Kapitel wurden letzten Endes auch Selektionsvorgänge beschrieben, die bewirken, daß "aus einem Wahrnehmungsgesamt Teile früher und klarer bewußt wahrgenommen werden, während andere Teile später weniger klar bewußt oder überhaupt nicht apperzipiert werden." (KLEBER, 1976, S.44).
KLEBER (a.a.O., S.43) führt einige klassische Begriffe auf, die spezielle Ausprägungen der Selektivität darstellen:

- Assimilation
  bezeichnet die Angleichung der Wahrnehmung aktueller Informationen an bereits gespeicherte wahrgenommene Informationen

- Akzentuierung
  bezeichnet die einzelnen Informationen zugeschriebene Relevanz und deren Auswirkung auf die Wahrnehmung

- Selektive Sensitivierung
  bezeichnet das Herabsetzen der Wahrnehmungsschwelle für Informationen, für die der Wahrnehmende sensitiviert wurde. Daran schließt sich eine akzentuierende Selektion an

- Auffassung
  bezeichnet die jedem Wahrnehmenden spezifisch eigene Kombination der genannten Selektionstendenzen

- Kausalitätswahrnehmung
  auch Kausalattribuierung; bezeichnet die bei der Wahrnehmung von sozialen Ereignissen auftretende Tendenz, diesen Ereignissen Ursachen zuzuschreiben. In der Forschung zur Leistungsmotivation (z.B. HECKHAUSEN,1974) wurde dieses Konzept verwendet zur Feststellung der Ursachen, auf die Erfolg und Versagen vorwiegend zurückgeführt werden.

Nach WEINER (1971) werden als Ursachen für Erfolg und Mißerfolg sowohl internale (Begabung, Anstrengung) als auch externale (Aufgabenschwierigkeit, Zufall) Determinanten wahrgenommen. Nach Untersuchungen von BECKMANN (1973) und MEYER/BUTZKAMM (1975) tendieren Lehrer bei Erfolg der Schüler eher dazu, internal zu attribuieren und sich so den Erfolg selbst zuzuschreiben, bei Mißerfolg externale Gründe zuzuschreiben.
Diese Zuschreibung erfolgt allerdings im Rahmen eines Vergleichs zwischen aktuellen und gespeicherten Informationen über die Reizperson einerseits und zwischen aktuellen Informationen über die Reizperson und dem Kontext (etwa: die Klasse) andererseits:

- haben fast alle Schüler der Klasse Mißerfolg, wird vorwiegend die Aufgabenschwierigkeit als Grund dafür attribuiert
- hat ein Schüler stabilen Erfolg, so wird eine einzelne gute Leistung der Begabung zugeschrieben
- zeigt ein Schüler dagegen variable Leistungen, dann wird ein Erfolg der Anstrengung zugeschrieben. In der Bewertung der einzelnen Determinanten scheinen Lehrer wenigstens im Grund- und Hauptschulbereich die Anstrengung der Begabung vorzuziehen. (HÖHN, 1967; HOFER, 1969).

Eine Erklärungsmöglichkeit für die Selektivität der Wahrnehmung bietet die "Hypothesen-Theorie" (s.IRLE, 1975,S.84 f). Wie wir aus dem Modell des Wahrnehmungsprozesses von WARR und KNAPPER (1968) ersehen, treffen auf der Informationsebene aktuelle und gespeicherte Informationen über die Reizperson zusammen. In die gespeicherten Informationen sind über Rückkoppelung schon frühere Urteile eingeflossen. Die Aufnahme der aktuellen Information wird von diesem Wissen geprägt, das der Urteiler über die Reizperson hat oder zu haben glaubt (Stereotyp, Erwartung), was BRUNER (1951) als Wahrnehmungskompromiss zwischen dem 'objektiven Wahrnehmungsgegenstand' und dem 'Erwartungspaket' des Beobachters bezeichnet. Aus diesem Erwartungspaket bildet der Beobachter mehr oder weniger konkrete Hypothesen, die unterschiedlich 'stark' sind. Je vielfältiger eine Hypothese im Erwartungspaket verankert ist und je häufiger sie in der Vergangenheit als bestätigt

erlebt wurde, desto 'stärker' ist die Hypothese. Die Hypothesenstärke hat folglich einen Einfluß auf den Wahrnehmungskompromiss in der Weise, daß bei starken Hypothesen bereits geringe Hinweisreize (aktuelle Informationen) des Wahrnehmungsgegenstandes genügen, um die Hypothese zu bestätigen, während bei schwachen Hypothesen massivere und häufigere Hinweisreize zur Bestätigung nötig sind.

Ein Beispiel von KLEBER (1978, S.598) mag diese Erklärung illustrieren und dieses Kapitel beschließen:
"Wenn zum Beispiel ein Lehrer die 'starke Hypothese' gebildet hat: 'Schüler A sei fleißig' - weil der Schüler pünktlich und sauber in der Schule erscheint, sein Vater in einer höheren Stellung ist und der Schüler bereits positiv aufgefallen ist, der Lehrer auch das Bedürfnis hat, über eine große Anzahl von fleißigen Schülern seinen persönlichen Erfolg nachzuweisen - ,dann genügt die Erledigung einiger gestellter Aufgaben durch den Schüler, diesen als fleißig und gegebenenfalls auch begabt wahrzunehmen. Bei einer schwachen Hypothese oder der entgegengesetzten Hypothese muß der Schüler sich erheblich mehr anstrengen, wenn der Lehrer das Gleiche wahrnehmen soll."

2.5. Implizite Persönlichkeitstheorien

2.5.1. Theoretische Einordnung

Wie wir gesehen haben, tritt in der sozialen Wahrnehmung eine Anzahl von Tendenzen auf, die die Urteilsvalidität wesentlich beeinträchtigen.

In der Darstellung dieser Urteilstendenzen besteht einmal die Möglichkeit, diese Tendenzen in loser Reihenfolge nebeneinander zu stellen, (z.B. bei BRUNER u. TAGIURI, 1954; GRAUMANN, 1960; WARR u. KNAPPER, 1968), oder aber sie nach statistischen Kennwerten zu ordnen (z.B. CRONBACH, 1955), so wie sie sich dem empirisch arbeitenden Wissenschaftler erschließen.
Die Nicht-Existenz eines theoretisch fundierten, möglichst noch empirisch abgesicherten Ordnungsschemas - falls wir von der Verlegenheitslösung einer Anordnung gemäß der in statistischen Lehrbüchern eingehaltenen Reihenfolge der Abhandlung von Kennwerten absehen - spiegelt die Tatsache wider, daß eine differenzierte allgemeinpsychologische Analyse der Urteilstendenzen nach bestimmten kognitiven, motivationalen und emotionalen Bedingungen oder eine sozialpsychologische Erforschung nach Formen und Zielen der Interaktion noch nicht ausgearbeitet ist. (ULICH/MERTENS, 1973, S.81).
LAUCKEN (1974) strebt zwar eine systematische Analyse des "Alltagswissens" an, das man .."im tagtäglichen Umgang mit Mitmenschen benutzt, um deren Verhalten zu beschreiben, zu erklären und zu antizipieren" (a.a.O.,S.18), zielt dabei aber auf eine Rahmentheorie der "Gesamtheit der naiv-psychologischen Vorstellungen zum 'Funktionieren' von Menschen" allgemein ab (a.a.O.,S.38). Dieser Intention folgend, bewegen sich seine Überlegungen auf einem so hohen Abstraktionsniveau, daß der an Stichproben gebundene Empiriker bei der Untersuchung objektbezogener, individueller oder gruppenspezifischer Inhalte Deduktionsprobleme hat, wenn er in seiner Untersuchung den Kontakt zu dieser übergreifenden Rahmentheorie nicht verlieren will; zum anderen hat der Leser solcher Untersuchungen mit spezifisch empirischem Inhalt Einordnungsprobleme.

Diese Probleme bestehen, solange nur theoretisch expliziert wird, daß naiv-psychologisches Wissen um Disposition und Prozesse bei Existenz einer Zielvorstellung zur Planung eines Handlungsentwurfs und schließlich zu einem Verhalten führt, und solange nicht hinreichend geklärt ist, welches Wissen bei welchen Zielvorstellungen zu einem wie beschaffenen Alltagsmenschen zu welchem Verhalten führen.

Wir werden also doch CRONBACH's (1955) Ordnung der Urteilstendenzen folgen müssen, wobei wir uns bewußt sein müssen, daß es sich hierbei um statistische Kennwerte handelt, deren Rechengrundlagen in Rating-Skalen eingezwängte Urteile darstellen. Unter "Urteil" wird hier also der vom Urteiler auf einem bi-polaren Merkmalskontinuum (Rating-Skala) markierte Skalenwert verstanden.
Wir unterscheiden also: (nach ULICH/MERTENS,1973,S.81ff)

1. Mittelwertstendenzen
   wie z.B. das A d a p t a t i o n s n i v e a u (HELSON,1948), das eigentlich aus der Psychophysik stammt und einen individuellen Bezugsrahmen für Urteile bezeichnet, der von vergangenen und gegenwärtigen Reizen abhängt,

   das K o n t r a s t p h ä n o m e n (MURRAY,1938), das die Beobachtung einer Gegenläufigkeit der Urteile eines Beurteilers bei der Selbst- und Fremdeinschätzung auf dem gleichen Merkmalskontinuum bezeichnet,

   den M i l d e - E f f e k t ('leniency'-Effekt, GUILFORD, 1954; 'generosity error', CRONBACH, 1960), unter dem man die kulturspezifische (BRANDSTÄTTER, 1970) Tendenz zu günstigeren Beurteilungen versteht,

   den ' e r r o r o f n e g a t i v e l e n i e n c y' (GUILFORD, 1954), der eine spezifische Urteilstendenz zu 'strengen' Urteilen beschreibt; er ist als Verschiebung des Mittelwertes der Urteile zum negativen Pol hin zu beobachten.

2. Streuungstendenzen

z.B. den "error of central tendency" (HOLLINGWORTH, 1910; zit.nach GUILFORD, 1954, S.323;).[1] Damit wird die Tendenz von Urteilern bezeichnet, bei der Fixierung von Urteilen auf Rating-Skalen nur einen Ausschnitt des zur Verfügung gestellten Merkmalskontinuums zu verwenden, der jedoch nicht zwingend mit dem Mittelbereich des Kontinuums zusammenfällt. Statistisch wird diese Tendenz als geringe Streuung der vom Urteiler gekennzeichneten Skalenwerte erfaßt; psychologisch wird dieses Phänomen als geringe Urteilsdifferenziertheit interpretiert.

Die Mittelwerts- und Streuungstendenzen abschließend, soll darauf hingewiesen werden, daß es sich trotz des gegenteiligen Sprachgebrauchs in allen Fällen nicht um einen "error" (zufälligen 'Fehler') im statistischen Sinn handelt, sondern um einen "bias" (systematischen 'Fehler').

3. Korrelationstendenzen

CRONBACH (1955) interpretiert die Mittelwerte der anhand von Skalen abgegebenen Urteile als Stereotyp, ihre Streuungen als Prägnanz des Stereotyps (ähnlich der 'Kognitiven Komplexität' von BIERI, 1955). Die Korrelationen der markierten Skalenwerte werden von ihm als Hinweise auf Analogie- und Kausalschlüsse zugeschriebener Eigenschaften aufeinander interpretiert.
Inhaltlich verweist uns diese Interpretation auf das weiter oben schon dargestellte Modell des Wahrnehmungsprozesses von WARR u. KNAPPER (1968) (s. Kapitel 'Soziale Wahrnehmung'), in dem das Verarbeitungszentrum der Ort von 'inference rules' ist, die Analogieschlüsse von einem beobachteten Merkmal auf ein nicht erfaßtes Merkmal erlauben.

---

[1] zur Berichtigung von ULICH/MERTENS (1973, S.83,166): HOLLINGWORTH, H.L.: Experimental studies in judgement, New York, Science Press, 1913; nicht 1910. Siehe auch: GUILFORD, (1954, S.262).

M e t h o d i s c h schließen sich hier CLINE (1964) und TODD u. RAPPOPORT (1969) an, die nicht mehr nur die Korrelationen zwischen einzelnen Skalen-Markierungen der Urteiler errechneten, sondern die Interkorrelationen aller von den Urteilern markierten Skalenwerte und die Faktorenstrukturen der Korrelationsmatrizen als 'Implizite Persönlichkeitstheorien' interpretierten, wodurch sich also diese Erklärungskonstrukte methodisch gesehen als zu den Korrelationstendenzen zugehörende ausweisen.
Gerade an dieser Stelle wird das allgemeine Dilemma des empirisch arbeitenden Psychologen klar: Auf der Suche nach Erklärung psychischer Zustände und Vorgänge ist er gezwungen, den Untersuchungsgegenstand greifbar und meßbar zu machen. Greifbar wird der Gegenstand dadurch, daß er ein Konstrukt über den Untersuchungsgegenstand errichtet, das seine theoretische Definition, gewissermaßen seine 'Wahrnehmung' des Untersuchungsgegenstandes darstellt. Solange er mit seinen Überlegungen im theoretischen Bereich dieses Konstrukts bleibt und theorie-immanent logisch argumentiert, stimmt das Konstrukt mit dem Untersuchungsgegenstand überein, da diese Übereinstimmung mit theorie-immanenten Begriffen ja nicht widerlegt werden kann.
Begibt er sich jedoch daran, den mit dem Konstrukt gemeinten Untersuchungsgegenstand auch noch meßbar zu machen, etwa durch Festlegung von Methoden der Datenerhebung und mathematisch-statistischer Verrechnung dieser Daten, so wird schon nicht mehr der gemeinte Gegenstand, sondern nur noch das Konstrukt gemessen. Die Meßbarkeit des Konstrukts aber wird wiederum vom zur Verfügung stehenden Repertoire an Erfassungs- und mathematisch-statistischen Methoden diktiert - das Konstrukt muß also, um es meßbar zu machen in Begriffen der statistischen und Erfassungsmethoden ausgedrückt werden.
Wie das Vokabular der Sprache aber nicht beliebig rasterbar ist (man suche z.B. für 10 durchaus mögliche Farbmischungen von grün und blau sprachliche Bezeichnungen), so ist auch das Vokabular der Erfassungsmethoden und der statistischen Methoden nicht beliebig abstufbar und jedem Konstrukt anzupassen. In unserem Fall beinhaltet das erhebungstechnische Methoden-Vokabular :"Rating-Skala", auf der mathematisch-

statistischen Seite "Mittelwert, Streuung, Korrelation,usw."

Allein durch die Anwendung einer bestimmten Methode auf einen psychologischen Gegenstandsbereich wird also eine psychologische Theorie (Konstrukt) über diesen Gegenstandsbereich impliziert. (s.dazu: GIGERENZER, 1978,S.110).

Diese Arbeit kann nun nicht der Ort sein für umfangreichere wissenschaftstheoretische Erörterungen der Methodenbedingtheit psychologischer Konstrukte oder der Gültigkeit von Konstrukten für den gemeinten Gegenstand, ein Thema, das jedoch noch einmal bei der Diskussion der Validität und Reliabilität der Meßergebnisse angeschnitten wird.

2.5.2. Der Begriff "Implizite Persönlichkeitstheorie"

Die Beobachtung des Phänomens, daß bei auf Rating-Skalen explizierten Urteilen über andere Menschen Korrelationen zwischen den vor Beurteilern markierten Skalenwerten auftreten, hat BRUNER u. TAGIURI (1954) zu folgender Erklärung dieses Sachverhaltes geführt: Annahmen des Urteilenden über Zusammenhänge von Eigenschaften beeinflussen die Beurteilung anderer Menschen. Das Gesamt dieser Annahmen bzw. der Korrelationstendenzen bezeichnen die genannten Autoren als 'implizite Persönlichkeitstheorie' des Urteilers ('naive implicit personality theory').

CRONBACH (1955) erweiterte den Begriff, indem er neben den genannten Korrelationstendenzen auch Mittelwerts- und Streuungstendenzen unter seiner 'lay theory of personality' faßte.
HAYS (1958), CLINE (1964) und TODD u. RAPPOPORT (1969) veränderten wie bereits oben erwähnt, den Inhalt des Begriffs vom Methodischen her dergestalt , daß sie die Faktorenstruktur der Interkorrelationsmatrix der zugrundeliegenden Eigenschaftsschätzungen als implizite Persönlichkeitstheorie bezeichneten.
ULICH/MERTENS (1973, S.102) definieren implizite Persönlichkeitstheorien als kognitive Bezugsysteme der zwischenmenschlichen Orientierung, die durch interaktive Lernprozesse erworben und modifiziert werden und die sich in sozialen Situationen als handlungsrelevante Erwartungshaltungen aktivieren; von der methodischen Definition her folgen sie den letztgenannten Autoren.
LAUCKEN (1974) spricht zwar auch den sozialen Aspekt an (a.a.O., S.37), untersucht aber ausschließlich das Wissen des Alltagsmenschen zur Erklärung fremden Verhaltens und verwendet daher den Begriff "implizite Persönlichkeitstheorie" im Sinne von BRUNER u. TAGIURI (1954) als Zuschreibung von "Dispositionskoppelungen" (a.a.O., S.171).

Die Wahl des Begriffs "implizite Persönlichkeitstheorie" für das beschriebene Korrelationsphänomen ist nicht besonders günstig, da mit den Begriffsteilen 'implizit' und '-theorie' Implikationen verbunden sind, die zu einer falschen Einschätzung des damit Gemeinten führen könnten.

Mit 'implizit' ist zunächst gemeint, daß "im Gegensatz zu explizit.. Theorienbildung über Kategorien und Messung nicht unabhängig, sondern in der Person des Raters zugleich erfolgen" (KLEITER, 1973, S.202).

Obwohl die implizite Persönlichkeitstheorie in der Literatur häufig als Urteils-'Fehler' bezeichnet wird, ist es unzulässig, sie von vornherein als falsch zu betrachten. Der Argumentation von HANKE, LOHMÜLLER, MANDL (1978,S.42), die 'Richtigkeit' von impliziten Persönlichkeitstheorien an der Übereinstimmung mit "Befunden instrumentell überlegener Meßinstrumente (z.B. Tests) festzustellen, können wir von unseren Überlegungen zur Merkmalsvalidität her (siehe Kapitel 'Problem der Eigenschaften') nicht folgen; mißt man eine implizite am Maßstab einer expliziten Persönlichkeitstheorie, so nimmt man nur einen Konstruktvergleich vor, bei dem die Konstrukte der expliziten Persönlichkeitstheorie wohl einen Vorsprung an Objektivität haben, nicht jedoch an Validität für den Untersuchungsgegenstand - falls man für die Kritik der impliziten Persönlichkeitstheorien überhaupt solche testtheoretischen Kriterien heranziehen sollte.

Der Begriffsteil "-theorie" soll wiederum nicht andeuten, daß es sich um eine wissenschaftliche Theorie als " ein durch Denken gewonnenes System der Verbindung von Tatsachen zu einem widerspruchslosen Zusammenhang von Gründen und Folgen" (DORSCH, 1963, S.330) handelt.Es wird vielmehr ein Wissen des Alltagsdiagnostikers damit bezeichnet, der dieses Wissen im Gegensatz zum Wissenschaftler nicht systematisch hinterfragt in bezug auf dessen Konstituierung, empirische Absicherung und formal-konstruktiven Aufbau, sondern ihm relativ unkritisch vertraut.
(s. LAUCKEN, 1974, S.20).

Andere Begriffe wie "Persönlichkeitskonstrukte" oder wohl eher "individuelle Persönlichkeitskonstrukte" (personal constructs) nach KELLY (1955) konnten sich in der deutschen Literatur nicht durchsetzen - wohl auch wegen der Betonung, die KELLY von seiner psychotherapeutischen Ausrichtung her auf 'individuell' legte, während sich die meisten Untersucher in Anlehnung an HASTORF/SCHNEIDER/POLEFKA (1970) auf implizite Persönlichkeitstheorien als Gruppenstereotypen konzentrierten.

An die Arbeiten der beiden letztgenannten Autoren knüpft sich als eine weitere Frage, ob dieser Begriff im Singular oder im Plural zu verwenden ist, d.h. ob es plausible theoretische Gründe oder Untersuchungen gibt, die für die Existenz nur einer oder aber mehrerer impliziter Persönlichkeitstheorien sprechen.

LAUCKEN (1974, S.24 ff) meldet selbst an dem Titel seines Buches "Naive Verhaltenstheorie" Bedenken an, indem er einräumt, daß es sich bei seiner Darstellung durchaus nur um mögliche, nämlich kultur- und zeitspezifische naive Verhaltenstheorien handeln könnte.
Um wieviel wahrscheinlicher wird die Vermutung der Existenz einer größeren Anzahl von impliziten Persönlichkeitstheorien (im folgenden IPTn), wenn man sich auf eine konkretere Ebene der Untersuchung begibt, das "Eigenschaftsimplikationsgefüge in t e i l k o n k r e t i s i e r t e m Zustand betrachtet, sei es in Form von subjektbezogenen Individualtheorien..., sei es in Form von auf Klassen von Untersuchungsobjekten eingeengten Spezifizierungen." (LAUCKEN, 1974, S.36).
Gerade diese letztere Form der Untersuchung einer konkreteren inhaltlichen Ebene ist angesprochen, wenn wir in dieser Arbeit von IPTn von Lehrern in bezug auf Schüler sprechen.
Betrachten wir die IPTn mit ULICH/MERTENS (1973,S.102) als in interaktiven Lernprozessen erworben, so könnte man erwarten, daß in ihnen gesellschaftsbezogene und individuelle Erfahrungen sichtbar werden (HASTORF et.al., 1970,S.46); es ist also denkbar, daß Beurteiler je nach ihrer Gruppenzugehörigkeit verschiedene IPTn aufweisen, die mit der Gruppenzugehörigkeit der Beurteilten variieren.

Von diesem Aspekt abgeleitet, könnte die Aktualisierung einer speziellen IPT unter anderem auch davon abhängen, in welcher Rolle sich Urteiler und Beurteilter gegenüberstehen.

> Ziehen wir als Beispiel Lehrer als Urteiler heran, so ist denkbar, daß ein Lehrer jeweils eine andere IPT aktualisiert, je nachdem, ob die Beurteilten Schüler sind oder seine eigenen Kinder; bei Lehrern könnte man ferner spezifische IPTn nach Schichtzugehörigkeit beurteilter Schüler oder Schultyp erwarten oder Kombinationen (z.B. IPT von Gymnasiallehrern bezüglich Grundschichtkindern, die das Gymnasium besuchen).
>
> Andererseits ist das Rollenverhältnis Lehrer-Schüler durch die grundsätzlichen Anforderungen der Gesellschaft und der Institution Schule gekennzeichnet, die diese dem Lehrer gegenüber stellen: Vermittlung von Wissen und gesellschaftliche Vorauslese nach dem Leistungskriterium. Aus diesem Aufgabenkreis ergeben sich für Lehrer verbindliche Maßstäbe der Leistungs- und Persönlichkeitsbeurteilung, die sich eventuell nivellierend auf die IPTn von Lehrern bezüglich Schülern auswirken könnten, was der anfangs geäußerten Vermutung sehr spezifischer IPTn widersprechen würde.
>
> So konnten z.B. HOFER (1969) und HANKE/LOHMÜLLER/MANDL (1978) in ihrer Lehrerstichprobe nur eine IPT feststellen, ebenso wie COHEN (1969) in einem ebenfalls begrenzten sozialen Raum (Militär) und LEVIN (et.al.,1973); COHEN u. SCHÜMER (1969); WALTERS u. JACKSON (1966); TODD u. RAPPOPORT (1969). HUBER/MANDL (1979) dagegen stellten in einer methodenkritischen Arbeit vier Beurteilungstypen von Lehrern fest.

Die Auswirkungen individueller Erfahrungen auf die IPTn könnte auf die ganz personspezifische soziale Lerngeschichte des Urteilers zurückzuführen sein. So könnten IPTn sozialer Aufsteiger durch eine besondere Konformität mit den in der Gruppe vorrangig verwendeten IPTn festgestellt werden, in die aufgestiegen wird. Der Grad der individuellen Identifikation mit der Rolle, die sich im individuell unterschiedlichen Spielen dieser Rolle ausdrückt, könnte also ebenfalls unterschiedliche IPTn determinieren.
JONES u. DE CHARMS (1958) fanden zu diesem Untersuchungsgegenstand, daß die durch die soziale Rolle festgelegte Art der Interaktion die Wahl der zur Beurteilung verwendeten Kategorien eher determiniert als die jeweiligen Persönlichkeitsmerkmale der Beurteiler, was die Annahme eines Dominierens rollenspezifischer IPTn über personspezifische IPTn nahelegt.

Wenn wir im folgenden wider besseres Wissen von impliziter Persönlichkeitstheorie im Singular sprechen, so ist das einmal darin begründet, daß die vorliegende Arbeit in Anlehnung an die Vorbildarbeit von HOFER (1969) schon 1971 geplant wurde mit dem Ziel eines Vergleichs der Ergebnisse, was ein Abweichen in den Methoden der Erhebung und statistisch-mathematischen Berechnung nicht ratsam erscheinen ließ. Wir sind uns klar darüber, daß wir durch das weitgehende Nachvollziehen der Methodik der Vorbildarbeit ebenfalls nur zur Feststellung einer IPT der Lehrergruppe kommen können. Die Fragestellungen der Arbeit intendieren jedoch kein Auffinden subgruppenspezifischer IPTn innerhalb der verwendeten Lehrerstichprobe.

2.5.3. Interpretationsmodi der Impliziten Persönlichkeitstheorien

In der Interpretation der IPTn wird deren Orientierungs- und Entlastungsfunktion hervorgehoben, (ULICH/MERTENS, 1973, S.102). Danach soll eine Reduktion der Vielzahl von Verhaltensdaten der Interaktionspartner erreicht und die Konstruktion eines sinnvollen, also interaktionsadäquaten Handlungsraumes ermöglicht werden.
Die grundlegende Interpretationsfrage bezieht sich nun darauf, ob die IPT "tatsächliche" Reizkonfigurationen des Beurteilten widerspiegelt (realistische Interpretation) oder ob in ihr nur kognizierte Vermutungen über Eigenschaftszusammenhänge zum Ausdruck kommen (idealistische Interpretation).
Für die Interpretation unter idealistischem Aspekt spricht theoretisch die auf die Aktivierung der IPT erfolgende subjektive Reduktion von Unsicherheit. Das subjektive Sicherheitsgefühl des Urteilers erhöht die Wahrscheinlichkeit der Anwendung der unveränderten IPT. Etwa im weiteren Verlauf der Interaktion auftretende IPT-dissonante Informationen müssten dann selegiert werden, d.h. der Wahrnehmungsgegenstand wird den Erwartungen angeglichen, womit kognitive Dissonanzen vermieden werden.

> Eine idealistische Interpretation legen Befunde von PASSINI/NORMAN (1966) nahe, in denen Faktorenstrukturen von Beurteilungen nach einer Kontaktzeit von 15 Minuten weitgehend denen von Beurteilungen aufgrund mehrjähriger Bekanntschaft entsprachen. Allerdings konnte in dem Urteilen über Bekannte eine größere Zahl signifikanter Faktoren extrahiert werden.

Für die Interpretation unter realistischem Aspekt könnte die mit der Interaktionshäufigkeit und -dauer steigende Wahrscheinlichkeit des Auftretens massiv IPT-dissonanter Informationen sprechen, die nun nicht mehr selegiert werden können, wenn der Urteiler das kognitive Gleichgewicht mit der sozialen Umwelt nicht verlieren will; wenn die IPT als naive Dispositionstheorie bei der Planung von Handlungsentwürfen als mitbeteiligt angesehen wird, würde eine hartnäckige Selektion massiv IPT-dissonanter Informationen die Planung interaktions·adäquater Handlungsentwürfe nach sich ziehen.

> Eine realistische Interpretation der IPT unterstützen die Ergebnisse von HOFER (1969), der beim Vergleich der Faktorenstrukturen aus Ähnlichkeitsschätzungen, Beurteilungen von Stereotypen und aktuellen Schülerbeurteilungen von Lehrern eine hohe Strukturkongruenz feststellte und mit diesem Ergebnis die hohe Ähnlichkeit der kognizierten und perzipierten Zusammenhangsvermutungen belegte.

Nicht zwingend für diesen Aspekt sprechen hingegen die Befunde, die HANKE/LOHMÖLLER/MANDL (1978, S.44) referieren. In ihren Untersuchungen wurde zwischen Grundschullehrerurteilen im vierten Schuljahr über die Gymnasialeignung ihrer Schüler und dem Gesamtwert aus verschiedenen objektiven Testverfahren eine Korrelation von .76 gefunden. Die in den letzten Jahren verstärkte Konfrontation der Lehrer in Ausbildung und Beruf mit Testverfahren könnte hier zu einer Angleichung der alltagspsychologischen an die wissenschaftlichen Konstrukte geführt haben, sodaß die von Testkonstruktoren für valide erachteten Kriterien von den Lehrern übernommen werden.
Als letzter Diskussionsbeitrag seien noch die in der IPT-Forschung verwendeten Methoden der Datenerhebung genannt, die aspektspezifischen Ergebnissen Vorschub leisten. Die Vorgabe von Beurteilungsitems auf Konstruktebene (Eigenschaften), die zudem oft nur vom Untersucher für relevant erachtet werden, produzieren wegen der Anfälligkeit von Konstrukten für Beurteilungstendenzen (hohe Inferenz; s. SCHWARZER/SCHWARZER, 1977, S.47 ff) eher Ergebnisse, die den idealistischen Interpretationsaspekt favorisieren; die Verwendung von Situationsverhaltens-Items, konkret beobachtbaren Verhaltensbeschreibungen also, senkt die Anfälligkeit der Rating-Listen für Beurteilungstendenzen; es handelt sich dann aber auch nicht mehr um Beurteilungen, sondern um Beobachtungen; der Untersucher muß nun den Legitimationsnachweis führen, wenn er diese Daten irgendwie zu kategorisieren versucht. SCHWARZER/SCHWARZER (1977,S.42ff) schlagen den Mittelweg eines mäßig inferenten Situations-Verhaltens-Ratings vor, bei dem Verhaltensweisen in standardisierten Situationen vom Urteiler eingeschätzt werden.

2.5.4. Methoden zur empirischen Erfassung von Impliziten Persönlichkeitstheorien

Empirisch zu überprüfende Konstrukte werden letztlich durch die Methoden der Erhebung der Daten und der statistisch-mathematischen Verrechnung dieser Daten definiert. Wie bereits im Kapitel "Theoretische Einordnung" angesprochen, hat das Konstrukt 'implizite Persönlichkeitstheorie' schon mehrere begriffliche Vorformen mit entsprechend unterschiedlichen erhebungs- und statistisch-methodischen Definitionen gehabt.

ASCH (1946) hat bereits mit seinen Experimenten zum "ersten Eindruck" Zusammenhänge zugeschriebener Eigenschaften festgestellt. Bei Veränderung einzelner vorgeschriebener Eigenschaften speziell aus dem emotional-affektiven Bereich (warmherzig, kühl, höflich, unhöflich) stellte er eine Veränderung der Zuschreibung fest; diese Eigenschaften bezeichnete er als 'zentrale Eigenschaften'.

KELLY (1955) ging es eigentlich nicht um die Feststellung von Urteilstendenzen, sondern um die individuelle Verwendung bestimmter 'personal constructs', die er klinisch-diagnostisch verwertete. Er ließ seine Vpn das Gegenteil von jeweils vorgegebenen Eigenschaftsbegriffen nennen und diagnostizierte aus diesen Nennungen Einstellungen der Vpn bezüglich dieses soeben skalierten Sachverhaltes. Wegen der vielen interpretatorischen Implikationen ist dieses Verfahren zumindest als nicht objektiv zu bezeichnen.

BRUNER, SHAPIRO u. TAGIURI (1958) beschrieben in ihrer Arbeit zur "naive implicit personality theory" die Ähnlichkeitsskalierung als Erhebungsmodus. Sie wird in zwei Formen praktiziert:

- Auf einer Rating-Skala soll beurteilt werden, wie groß die Ähnlichkeit zweier vorgegebener Eigenschaften ist.
- Auf einer Rating-Skala soll beurteilt werden, wie wahrscheinlich es ist, daß ein Mensch mit einer Eigenschaft auch die andere Eigenschaft aufweist oder in welchen Ausprägungsgraden beide Eigenschaften bei einem Menschen vorkommen.

Die erstgenannte Form hat eher eine Überprüfung der Synonymität und semantischen Kongruenz zum Ziel, während die zweite Form dem Konstrukt eher angemessen erscheint, da sie konkreter nach Vermutungen über tatsächliche Eigenschaftszusammenhänge fragt.

Die Trennung mag rein artifiziell anmuten, da sich die Vpn mit großer Wahrscheinlichkeit auch bei der ersten Form konkrete Personen vorstellen, um die Aufgabe bewältigen zu können; HOFER (1969) vermischt beide Formen.

CRONBACH (1955) fügt in seiner 'lay theory of personality' der Untersuchung von Korrelationstendenzen noch Mittelwerts- und Streuungstendenzen hinzu und verstand unter 'impliziter Persönlichkeitstheorie' das Gesamt der nach statistischen Kennwerten geordneten Urteilstendenzen.

Mit HAYS (1958) trat das Konstrukt der IPT in seine moderne methodische Form. Er verwendete das Erhebungsverfahren von BRUNER et.al. (1958), unterzog die Interkorrelationen jedoch einer Faktorenanalyse und bezeichnete die aufgefundene Faktorenstruktur als IPT.

KEIL (1970) verwendete zur Ähnlichkeitsschätzung ein graphisches Distanzrating, bei dem der von den Vpn eingestellte Winkel zwischen zwei Zeigern den Grad der Ähnlichkeit der Eigenschaften angeben sollte.

ROSENBERG et.al. (1968) gingen bei der Datenerhebung mit der Methode des 'trait-sorting' vor, nach der der Urteiler Personen seiner Wahl beschreiben soll, indem er aus vorgegebenen Listen Eigenschaftsbezeichnungen aussucht, wobei jede

Eigenschaft nur einmal verwendet werden darf.

HUBER/MANDL (1979) äußern Kritik einmal an der Erhebungsmethode; sie ersetzen die üblicherweise vorgegebenen Listen von Rating-Skalen mit Eigenschaftsbezeichnungen durch nicht an die Verwendung von Eigenschaftsbegriffen und Quantifizierungen gebundene 'naturalistische' Persönlichkeitsbeschreibungen (a.a.O., S.5). Um die so provozierte Vielfalt des Materials zu reduzieren und damit einer statistisch - mathematischen Analyse zugänglich zu machen, sind die Autoren dann allerdings doch gezwungen, in einer Inhaltsanalyse Manipulationen vorzunehmen wie Bedeutungsinterpretationen und Kategorisierungen, womit die anfängliche 'Lebensnähe' der erhobenen Daten (GIGERENZER u. STRUBE, 1978, S.85) zumindest stark reduziert wird.

Weitere methodische Veränderungsvorschläge, die speziell von GIGERENZER u. STRUBE (1978) gemacht wurden, werden wegen des thematischen Zusammenhangs im Anschluß an die Darstellung der Vorbildarbeit von HOFER (1969, 1974) angesprochen.

Die weitere statistisch - mathematische Verrechnung der Daten mit dem Ziel der Erstellung eines Modells der Urteilsstruktur basiert auf der Matrix Eigenschaften x Eigenschaften. Werden Korrelationen errechnet, so werden zur Modell-Erstellung faktorenanalytische Verfahren angewandt (PAWLIK, 1967); werden die erhaltenen Maße als Distanzen aufgefaßt, kommen in der Folge Verfahren der multidimensionalen Skalierung zur Anwendung (KÜHN, 1976). Ein Beispiel für beide Vorgehensweisen und ein Vergleich der Ergebnisse findet sich bei HOFER (1969).

Um ein Modell des hierarchischen Aufbaus der verwendeten Eigenschaften zu erhalten, werden entweder Faktorenanalysen höherer Ordnung durchgeführt (PAWLIK, 1968, S.247 ff.), oder es wird die Technik der hierarchischen Clusteranalyse (ROLLETT u. BARTRAM, 1976) angewandt.

III Fragestellung

3.1. Begleitumstände und Ziele der Arbeit und ihre Auswirkungen auf Fragestellung und Methodik.

Wie schon einleitend bemerkt, entstanden Plan und statistisches Design zur vorliegenden Untersuchung in den Jahren 1971/72 unter dem Eindruck der Arbeit von HOFER (1969), die auch die Vorbildarbeit darstellt. Die Erhebung der Daten wurde 1973/74 durchgeführt, sodaß die methodenkritische Literatur, die z.T. sehr viel später erschien, nur noch in Form von Erklärungen oder Darstellungen eingearbeitet wurde und sich nicht mehr formativ auf die Arbeit auswirken konnte.

Besonders traf das für die mit dieser Arbeit verbundenen Probleme der Veränderungsmessung zu, da die weitverstreute, hauptsächlich im englischsprachlichen Raum erscheinende Literatur zu diesem Thema nicht nur dem Autor schwer zugänglich war. HELMREICH (1977) hat eine erste zusammenfassende Darstellung und Ordnung der Literatur zu diesem Themenkreis versucht, gefolgt von PETERMANN (1978), allerdings zu einem Zeitpunkt, zu dem die vorliegende Arbeit bereits in weiten Teilen fertiggestellt war.
Die wiederholte Überarbeitung konnte zwar wohl krassere Mißstände beseitigen, es besteht jedoch nach wie vor die Gefahr, daß Fragestellung und Methodik veraltet erscheinen - besonders dem Leser, der über ein dem heutigen Stand der Diskussion angepaßtes Methodenarsenal verfügt.

Unter drei Aspekten wird diese Ansicht des Lesers jedoch relativiert:

- Einmal war und ist die vorliegende Arbeit als Korrespondenzarbeit zur Untersuchung von HOFER (1969) gedacht, mit dem Ziel des Vergleiches der an zwei verschiedenen Stichproben gewonnenen Ergebnisse. Von dieser Zielformulierung her ist eine Angleichung der Fragestellung und Methodik ratsam (siehe dazu: COHEN, 1969, S.6).

- Zum anderen ist die Erweiterung der Fragestellung auf die Variabilität der IPT über eine bestimmte Zeit hinweg auch im Jahr 1979 noch - sieht man einmal von der Arbeit von HANKE, LOHMÜLLER und MANDL (1978) ab - in der Literatur zur impliziten Persönlichkeitstheorie nicht allzu häufig zu finden.

- Betrachtet man schließlich als Gradmesser des Interesses an der grundlegenden Fragestellung die (teils erweiterten) Neuauflagen der Originalarbeit von HOFER (1969) - die dritte Auflage erschien im Jahr 1974-, ferner die Zeit-Punkte der Veröffentlichung von Replikationen mit gleicher Methode der Datenerhebung : MASENDORF,TSCHERNER und TÜCKE (1974); TSCHERNER und MASENDORF (1974) und mit zusätzlich gleicher statistischer Methode : SCHWARZER (1976), so wird die Ansicht einer antiquierten Fragestellung doch wenigstens etwas relativiert.

3.2. Vorbild-Arbeit

Als Vorbild dient die Arbeit von HOFER: "Die Schülerpersönlichkeit im Urteil des Lehrers" (1969), deren Zusammenfassung im folgenden zitiert werden soll.

Fragestellung

"1. Welche "implizite Persönlichkeitstheorie", d.h. welche vorgefaßte Theorie über die Persönlichkeitsdimensionen von Schülern besitzen Lehrer?
2. Wie groß ist die Homogenität der Sichtweisen bei Lehrern hinsichtlich dieser Theorie?
3. Welche Persönlichkeitseigenschaften determinieren die Art und Weise, in der Lehrer die Schülerpersönlichkeit wahrnehmen?
4. Welchen Einfluß auf die aktuelle Schülerbeurteilung besitzt die implizite Persönlichkeitstheorie der Lehrer?
5. Führen verschiedene Methoden zur Erfassung der impliziten Persönlichkeitstheorie von Lehrern über Schüler zu demselben Ergebnis?"

Methode

"Eine Gruppe von 40 Lehrern wurde gebeten, 25 Persönlichkeitseigenschaften von Schülern (z.B. aufgeweckt, fleißig, gesellig, schüchtern, unordentlich) anhand einer 9-Punkte-Skala paarweise nach Ähnlichkeit bei Schülern zu beurteilen. Die Matrix der Ähnlichkeitsschätzungen wurde (mit Hilfe der Tucker- und Messick - Methode) daraufhin untersucht, ob die Eigenschaftsrelationen von allen Lehrern in gleicher Weise gesehen werden. Die resultierenden Sichtweisen (Beurteilerfaktoren) wurden dann, jede für sich, multidimensional analysiert. Auf diese Weise konnten die verschiedenen impliziten Persönlichkeitstheorien ermittelt werden. Die Ladungen der einzelnen Lehrer auf diesen Beurteilerfaktoren wurden dann mit 13 an ihnen erhobenen Persönlichkeits- und Einstellungsmessungen (Leistungsmotivation, Lehrereinstellungen zu Kind und Unterricht, Autoritarismus, direktive Einstellung, Höflichkeit, Extraversion, Neurotizismus, Rigidität, Differenz zwischen Selbst- und Wunschbild, Alter, Zahl der Schüler in der Klasse, Schuljahr, Länge der Schulpraxis) korreliert.
Eine zweite Gruppe von 66 Lehrern wurde gebeten, fünf jeweils streng nach Zufall ausgewählte Schüler ihrer Klasse auf den gleichen, zu 7-stufigen Skalen erweiterten Eigenschaften zu beurteilen. Die Faktorenanalyse dieser Beurteilungen ergab die Faktorenstruktur aktuell beurteilter Schüler. Sie wurde mit der durch multidimensionale Skalierung der Ähnlichkeitsschätzungen ermittelten Faktorenstruktur analytisch verglichen.

Diese zweite Gruppe von Lehrern hatte außerdem vier Schülertypen (den durchschnittlichen, guten und schlechten Schüler ihrer Erfahrung sowie den idealen Schüler) auf gleichen Eigenschaftsskalen zu beurteilen. Die Beurteilungen wurden für jeden Lehrer pro Skala gemittelt. Die Faktorenanalyse der Mittelwerte sollte ebenfalls die Struktur der impliziten Persönlichkeitstheorie widerspiegeln. Diese Struktur wurde dann mit der Struktur der aktuell beurteilten Schüler verglichen. Außerdem wurden an diesen 66 Lehrern dieselben 13 Persönlichkeits- und Einstellungsmessungen wie an den Lehrern der ersten Stichprobe erhoben und mit den Faktorenscores der Lehrer auf den Faktoren der aktuell beurteilten Schüler korreliert."

Ergebnisse und Interpretation

"1. Die implizite Theorie von Lehrern im Hinblick auf die Schülerpersönlichkeit besitzt im Rahmen der vorgegebenen 25 Eigenschaften fünf Dimensionen. Diese fünf der multidimensionalen Skalierung der Eigenschaftsähnlichkeiten entnommenen Faktoren wurden folgendermaßen benannt:
I. Arbeitsverhalten (mit den Eigenschaften 'konzentriert', 'pflichtbewußt', 'ordentlich');
II. Schwierigkeit ('schüchtern', 'sensibel', 'kompliziert');
III. Begabung ('intelligent', 'begabt', 'einfallsreich');
IV. Dominanz ('geltungsbedürftig', 'ehrgeizig'); und
V. soziale Zurückgezogenheit ('verschlossen','ungesellig').

Diese Faktoren wurden dahingehend interpretiert, daß sie das subjektive Ordnungsschema von Lehrern darstellen, das es ihnen ermöglicht, die Schülerpersönlichkeit schnell und einfach, d.h. mit wenigen Informationen, zu erfassen.
2. Die Lehrer sehen die Ähnlichkeiten zwischen den Eigenschaften in übereinstimmender Weise, denn es resultierte nur ein einziger Beurteilerfaktor. Diese Homogenität der Beurteilungen wird auf ein weitgehend gemeinsames System wahrgenommener Eigenschaftsbeziehungen zurückgeführt. Allerdings gibt es zwischen den Lehrern einige systematische Unterschiede, denen näher nachzugehen versucht wurde.
3. Es ergaben sich schwache Zusammenhänge zwischen den 13 Variablen und den Ausprägungsgraden der Lehrer auf dem einzigen Beurteilerfaktor. Und zwar zeigte sich lediglich, daß 'psychisch stabile', 'rigide' und ältere Lehrer eine schematische Beurteilung bevorzugt praktizierten im Gegensatz zu 'psychisch labilen', 'flexiblen' und jüngeren Lehrern. Dieses Ergebnis wurde dahingehend interpretiert, daß die genannten Variablen (psychisch stabil, rigide, alt) es erleichtern, allgemeine und vorgefaßte Meinungsstrukturen zu übernehmen.
4. Die Übereinstimmung zwischen der Faktorenstruktur der impliziten Persönlichkeitstheorie und jener der aktuell beurteilten Schüler war erstaunlich hoch: die Übereinstimmungskoeffizienten der einzelnen Faktoren lagen zwischen .86 und .96, jene der einzelnen Eigenschaften bzw. Eigenschaftsskalen zwischen .74 und .99. Dies wurde dahingehend interpretiert, daß die Beurteilung bekannter Schüler durch ihre Lehrer weniger durch die Persönlichkeit der Schüler als vielmehr durch die implizite Persönlichkeitstheorie der Lehrer bestimmt wird.

5. Die beiden Methoden zur Erfassung der impliziten Persönlichkeitstheorie, die MDS von Ähnlichkeiten zwischen Eigenschaften und die FA von subjektiven Erwartungswerten im Hinblick auf Schülertypen, führten zu einer weitgehenden Äquivalenz der Strukturen, wenngleich die Ähnlichkeitskoeffizienten wegen gewisser Insuffizienzen der zweiten Methode nicht in der erwarteten Höhe lagen. Die FA der subjektiven Erwartungswerte wurde daher als für weniger gut zur Erfassung der impliziten Persönlichkeitstheorie geeignet angesehen als die MDS von Ähnlichkeiten zwischen Eigenschaften."
(HOFER, 1969, S. 95 ff.) (Unterstreichungen vom Verfasser.)

3.3. Eingrenzung der Fragestellung

COHEN (1969, S.6) möchte in empirischen Untersuchungen systematischer Tendenzen bei Persönlichkeitsbeurteilungen neben anderen auch die folgende Auflage berücksichtigt wissen:" Es genügt nicht, empirisches Datenmaterial nur unter dem Gesichtspunkt einer ganz speziellen Hypothese zu überprüfen. Beim gegenwärtigen Stand der Theorienbildung zu den Phänomenen zwischenmenschlicher Beurteilung kann eine empirische Untersuchung nur dann als Beitrag für eine spätere Synthese verschiedener Fragestellungen dienen, wenn das jeweilige Untersuchungsmaterial soweit als möglich auch unter Fragestellungen analysiert wird, die im Mittelpunkt des Interesses anderer Autoren standen, welche mit ähnlichem Datenmaterial gearbeitet haben."
Von diesen Überlegungen her ist die teilweise starke inhaltliche und methodische Ähnlichkeit mit der Vorbildarbeit zu verstehen.
Noch eine letzte Bemerkung bezüglich der mit den Fragestellungen verbundenen Explikationen sei hier eingefügt.

GIGERENZER u. STRUBE (1978) legen am Beispiel der Arbeit von HOFER (1969) ihre Auffassung der Explizierung von Annahmen dar. Danach sollen alle Annahmen inhaltlicher und methodischer Art, speziell auch die "allgemein bekannten" über den ganzen Ablauf der Untersuchung hinweg permanent expliziert werden. Die Autoren gehen von der Überlegung aus, daß bei einer Annahme, die nicht durch eine genaue Explikation auf ihre Gültigkeit hin untersucht wurde, die Artefakt-Vermutung den Charakter einer Alternativ-Hypothese zur Forschungshypothese gewinnt. Sie fordern daher eine erschöpfende Explizierung jeder Annahme alternativ durch theoretische Begründung, empirische Überprüfung oder Reduktion (Verzicht auf eine Annahme zugunsten einer abgeschwächten Annahme).
Da sich die theoretische Begründung auf eine empirisch gestützte Theorie beziehen sollte - und solche Theorien sind in der Psychologie selten - drehen wir uns hier im Kreis . Folglich bleibt für die Explikation nur das Verfahren der Reduktion. Da wir aber wegen des angestrebten Vergleichs mit der Vorbildarbeit nur wenige inhaltliche und keine methodischen Reduktionen (z.B. Verwendung nicht-parametrischer Verfahren) vornehmen können, werden wir mit Sicherheit dem von den Autoren geforderten Konzept der 'annahmekritischen Dimensionsanalyse'nicht genügen können.

Diesen Vorbemerkungen schließt sich die erste Frage an:

1. Welche implizite Persönlichkeitstheorie haben die Lehrer der untersuchten Stichprobe in bezug auf Schüler?

In Kapitel 2 wurde durch Darstellung von Theorien und Untersuchungen die Vermutung begründet, daß Wahrnehmung und Beurteilung von einer naiven Theorie über Eigenschaftszusammenhänge beeinflußt werden.
Das Konstrukt der 'impliziten Persönlichkeitstheorie' (IPT) soll diesen vermuteten Sachverhalt modellhaft abbilden. Dazu wird auch in dieser Arbeit die statistische Methode der Faktorenanalyse verwendet. Ausgangsdaten sollen Skalenpunkte sein, durch deren Markierung Urteiler ihr Urteil über fiktive Personen zum Ausdruck bringen sollen. Bei den verwendeten Skalen handelt es sich um Rating-Skalen; die beiden Pole der Skalen sind mit Eigenschaftsbegriffen versehen, die von den Urteilern erfragt wurden, da sie von den Urteilern als relevant für die Beurteilungsobjekte in der Beurteilungssituation erachtet werden sollen. Die Faktorenstruktur der Korrelationsmatrix der Eigenschaften soll unser Modell (IPT) sein, von dem wir annehmen, daß es die naive Theorie unserer Urteiler über Eigenschaftszusammenhänge wiederspiegelt.

Da die Trennung beider Begriffe recht umständliche Formulierungen notwendig macht, und zudem der Nachweis der Gültigkeit des Konstrukts IPT für den eigentlich gemeinten Sachverhalt "naive Theorie über Eigenschaftszusammenhänge" nicht erbracht werden kann, werden wir nur noch von 'impliziter Persönlichkeitstheorie' sprechen, ihren Modellcharakter jedoch dabei berücksichtigen.
Da wir ferner aus der Literaturdiskussion die Vermutung ableiten, es könne sich um gruppenspezifische naive Theorien handeln, wurde die Urteiler-Gruppe so homogen wie möglich gewählt, unter Berücksichtigung der sozialen Rollen (Lehrer-Lehrerinnen), des Beschäftigungsverhältnisses (Beamte), der Art der Beschäftigung (Lehrtätigkeit in der ersten Klasse eines Gymnasiums) und des Beschäftigungsortes (Gymnasium in einer Stadt).

Die Homogenität der Stichprobe schließt allerdings eine Generalisierung der Ergebnisse aus. Die Ergebnisse betreffen also nur diese eine Stichprobe.
HOFER (1969) versucht, den Eindruck einer Generalisierbarkeit seiner Ergebnisse zu retten - indem er eine gewisse Repräsentativität seiner Urteilerstichprobe andeutet, dabei aber weder das Verfahren der Stichprobenziehung noch die Population näher beschreibt:"Wie die Zahlen zeigen, handelt es sich aller Voraussicht nach um eine Stichprobe, die in den angegebenen Variablen die gesamte Breite der möglichen Population annähernd repräsentiert und die in Bezug auf die angesprochenen Variablen recht inhomogen ist." (a.a.O., S.41).
Es soll in dieser Studie nicht der Eindruck von Repräsentativität erweckt werden. Der geringe, vom Untersucher aber gerade noch organisatorisch zu bewältigende Stichprobenumfang und die Einschränkung auf einen eng umgrenzten Untersuchungsraum lassen das nicht zu. Diese Einschränkung erfüllt zwar die eine Forderung COHENs (1969): "Die Untersuchung sollte sich auf einen einheitlichen und überschaubaren ökologischen Bereich beschränken, sowohl bezüglich der Beurteiler als auch bezüglich der beurteilten Personen und der jeweiligen Aufgabenstellung." (a.a.O., S.6), macht es aber unmöglich, in dem von uns erfassten ökologischen Bereich eine zweite Stichprobe zu ziehen um mittels eines Kreuzvergleiches die Übertragbarkeit der Ergebnisse zu prüfen. Wäre die Stichprobe geteilt worden, so hätten zwar zwei annähernd gleiche Stichproben vorgelegen, es hätte dann jedoch vom Stichprobenumfang her keine konzeptadaequate statistische Bearbeitung der Daten erfolgen können. Wäre andererseits der Begriff des 'einheitlichen ökologischen Lebensraumes' nur um ein weiteres Kriterium für die Zugehörigkeit zur Stichprobe erweitert worden, (z.B. durch Einbeziehung eines weiteren Ortes oder der Lehrer einer weiteren Klassenstufe), so hätte diese Stichprobe nicht mehr von einem einzigen Untersucher bewältigt werden können.

Um Daten für die IPT der Lehrer in Bezug auf "Schüler" zu gewinnen, wurden anhand der weiter oben angesprochenen Rating-Skalen Urteile über die Stereotypen "der schlechte Schüler", "der gute Schüler" und "der" Schüler eingeholt.

Als Antwort auf diese Fragestellung können wir unter den gemachten Voraussetzungen analog zur Vorbildarbeit mit dem Resultieren einer impliziten Persönlichkeitstheorie der Lehrer rechnen, die die Stichprobe bildeten.

Von einer IPT können wir aber wohl erst dann sinnvoll sprechen, wenn sich diese kognitiv vermuteten Eigenschaftszusammenhänge auch als relevant für die Beurteilung realer Personen aufzeigen lassen. Um darüber eine Aussage machen zu können, ob die bei der Analyse dieser 'subjektiven Erwartungswerte' (COHEN, 1969, S.158) festgestellten Faktoren auch in aktuellen Schülerbeurteilungen nachzuweisen sind, muß die nächste Frage lauten:

2. Wie ist die Urteilsstruktur derselben Lehrerstichprobe beschaffen, die aus aktuellen Schülerbeurteilungen extrahiert werden kann?

In der Bearbeitung dieser Frage weichen wir von der Vorbildarbeit ab:

HOFER (1969) verwendet für die Feststellung der Struktur von aktuellen Urteilen eine andere Stichprobe als die, anhand derer er das IPT-Modell entworfen hat. Da uns der gleichzeitige Wechsel der Beurteilungsobjekte und der Stichprobe im Hinblick auf die Notwendigkeit einer Interpretation eventueller Strukturinkongruenzen zu gewagt erscheint, werden wir die Stichprobe beibehalten.

Ein weiterer Unterschied zwischen den vorliegenden und der Vorbildarbeit besteht darin, daß dort der Bekanntheitsgrad zwischen Urteilern und Beurteilten nicht näher beschrieben wird, die wir jedoch genauer definieren wollen.
PASSINI u. NORMAN (1966), die aus Urteilen über relativ fremde Personen (15 Minuten Sichtkontakt ohne verbale Kommunikation) die gleiche Urteilsstruktur darstellen konnten wie aus Urteilen über Personen, die die Urteiler z.T. schon

jahrelang kannten.bilden einen Ausgangspunkt unserer Überlegung.

CRONBACH (1958) und HAYS (1958) nehmen an, daß bei geringer Detailkenntnis in der Urteilssituation der Einfluß stereotyper Erwartungen verschiedener Eigenschaften besonders groß ist.
Setzen wir hierzu die Ergebnisse KOLTUVs (1962) in Beziehung, die hohe Interkorrelationen von Eigenschaftsbeurteilungen bezüglich fremder Personen, geringe Korrelationen jedoch bei der Beurteilung Bekannter fand, so ist anzunehmen, daß die Faktorenstruktur aus den Stereotypen-Beurteilungen in den Beurteilungen realer aber noch nicht gut bekannter Schüler besonders gut zu reproduzieren sein wird.

Die nächste Frage muß also lauten:

3. Sind die Faktoren aus der Beurteilung von Schülerstereotypen auch aus Beurteilungen realer aber noch relativ unbekannter Schüler zu reproduzieren?

Als "relativ unbekannt" sollen Schüler einer neuen Klasse bezeichnet werden, mit der die Lehrer eine Stunde Kontakt hatten und die anhand der gleichen Ratingskalen-Listen beurteilt werden sollen, die schon zur Beurteilung der Schülerstereotypen verwendet worden waren.

Die Frage mag wegen ihres schon vorhersehbaren Ergebnisses trivial anmuten, steht aber völlig im Einklang mit der Untersuchung von PASSINI u. NORMAN (1966), die, wie bereits dargestellt, bereits nach einem Kontakt von 15 Minuten eine Beurteilung durchführen liessen; diese Untersuchung dient in der Literatur stets als Beleg für den idealistischen Interpretationsmodus der IPT, soll aber in der Diskussion der nächsten Frage noch einmal angesprochen werden.

Anknüpfend an die Befunde KOLTUVs (1962) wäre mit steigendem Bekanntheitsgrad zwischen Beurteiler und Beurteilten eine immer geringere Reproduzierbarkeit der Faktorenstruktur zu vermuten. Nach weiteren Ergebnissen der Untersuchung von PASSINI und NORMAN (1966), die aus den Beurteilungen relativ unbekannter Personen nur 4 signifikante Faktoren extrahieren konnten, aus den Beurteilungen bekannter Personen dagegen

5 Faktoren, ließe sich ableiten, daß im Verlauf des Bekannt-Werdens von Urteilern und Beurteilten die Beurteilungen differenzierter werden. Man könnte annehmen, daß einzelne Urteilsdimensionen im Lauf des Kontakts an Wichtigkeit verlieren, andere gewinnen, was man aus dem Positionswechsel von Faktoren schließen könnte, der in einer Untersuchung von HANKE , LOHMÖLLER u. MANDL (1978, S.54) aufgetreten ist.

In dieser Hinsicht wollen wir bei der Untersuchung auch die 'zentrale Eigenschaft' in den Urteilen beachten. ASCH (1946) hat diesen Begriff eigentlich geprägt für Eigenschaftsbezeichnungen die den ganzheitlichen Eindruck affektiver Art von einer Person beeinflussen, der sich ergibt, wenn man eine Reihe von Eigenschaftsbezeichnungen liest, die diese Person beschreiben sollen. Aber wenn ASCH (1946) die "Psychologie des Gutachten-Lesens" erforscht hat (KAMINSKI, 1959, S.141), dann erforschen wir die 'Psychologie des Gutachten-Schreibens'.Wir wollen aber den Begriff nicht so eng fassen, sondern mit WISHNER (1960) als diejenige Eigenschaft (Skala), die die meisten und höchsten Korrelationen mit anderen Eigenschaften aufweist, aus faktorenanalytischer Sicht also die Skala mit der höchsten Kommunalität. Dabei wollen wir ~~aber~~ die emotionale Bewertung der wahrgenommenen Persönlichkeitsfaktoren nicht ganz außer Acht lassen. Um die affektiven Urteilsreaktionen ebenfalls zu erfassen, werden einzelne Skalen eingeführt, (z.B. "sympathisch", "hübsch"), die in den Nennungen relevanter Eigenschaften durch die Lehrer mit Sicherheit nicht auftauchen, da eine Beurteilung nach diesen Kategorien sozial nicht erwünscht ist.

4. Bleiben Inhalt und Bedeutung der Faktoren über die Zeit hinweg konstant oder verändern sie sich?

Bei der Beantwortung dieser Frage ist der Zeitfaktor unter den Voraussetzungen einer Feldstudie schwer zu erfassen. Da wir die Wahrnehmungstätigkeit des Lehrers weder beobachten noch messen oder gar kontrollierend auf sie einwirken können, müssen wir von der Fiktion einer Gleichverteilung der Chancen der Schüler ausgehen, vom Lehrer wahrgenommen

zu werden. Erst von seinem Verhalten schließen wir auf die vorangegangene Wahrnehmung. Dieses Verhalten, von dem aus wir rückschließen, daß Wahrnehmung stattgefunden hat, ist sein 'Ankreuz-Verhalten' auf den Ratingskalen-Listen, dessen Produkt wir auswerten und das natürlich recht wenig mit dem Ablauf seiner Alltagsdiagnostik zu tun hat.

Wegen dieser Unmöglichkeit der Kontrolle der Häufigkeit, mit der einzelne Schüler wahrgenommen werden, müssen wir auf den bloßen Ablauf gleichbleibender Zeitintervalle als Ordnungssystem für die auszuwertenden Informationen zurückgreifen.
Als Kriterium für Konstanz und Veränderung der Urteilsdimensionen sollen die im nächsten Teil der Arbeit an Ort und Stelle erläuterten statistischen Maßzahlen dienen. Unter Konstanz soll hier grundsätzlich die Identität der Anzahl, der Extraktionsposition und des Inhalts der Faktoren verstanden werden, unter Veränderung eine Abweichung in einem dieser drei Kriterien; die Darstellung und Interpretation dieser Abweichung soll die Art der Veränderung definieren.

Fundierte Hypothesen können wir wegen eines Mangels an vergleichbaren Untersuchungen nicht erstellen. Eine Untersuchung von HUBER, LOHMÖLLER und MANDL (1978) erbrachte keine Veränderung der Faktorenstruktur, wohl aber einen Wechsel des Extraktionsranges von Faktoren nachgeordneter Wichtigkeit.

Die einzige definitive Hypothese betrifft die aktuellen Beurteilungen zum ersten Meßzeitpunkt: die resultierenden Faktoren werden den Faktoren aus der Stereotypen-Beurteilung sehr ähnlich sein. Die Urteiler werden zu diesem Zeitpunkt, zu dem sie die Beurteilten noch kaum kennen, sicher schon über einige Beobachtungen einzelner Schüler verfügen, an denen sie ihre Vermutungen über Eigenschaftszusammenhänge festmachen können. Durch den Zwang, ein Urteil abgeben zu müssen, werden die Urteiler sich mit großer Wahrscheinlichkeit auf dieselben Ähnlichkeitsschätzungen der vorgegebenen Eigenschaften zurückziehen, die schon den Stereotypenbeurteilungen zugrundelagen. Da die Urteiler noch recht unsicher sind,

werden ihre Markierungen sicher eng um den O-Punkt der Skalen streuen, der dadurch seinen Charakter als mittlerer Ausprägungsgrad "weder-noch" der Beurteilungskategorie verliert und als "Indifferenzpunkt" die Unsicherheit der Urteiler widerspiegelt. Der Möglichkeit einer Feststellung der Faktorenstruktur tut diese Tatsache jedoch keinen Abbruch, da es hierfür nicht auf die markierten Ausprägungsgrade ankommt, sondern auf die zahlenmäßige Übereinstimmung der Ausprägungsgrade, die durch die Benützung eines eingeengten Skalenbereichs eher noch gefördert wird. Um den Zusammenhang dieser Zentraltendenz mit der Urteilssicherheit der Vpn zu überprüfen, wird eine Skala "subjektive Urteilssicherheit" eingeführt, auf der jeder Urteiler eine Sicherheitsschätzung des soeben abgegebenen Urteils vornimmt. Zu erwarten ist ein Nachlassen der Zentraltendenz mit steigender Urteilssicherheit, womit die Skala erst allmählich ihren Charakter als Rating-Skala der Eigenschaftsbegriffe erhält. Schon von dieser Überlegung her ist eine Veränderung der extrahierten Faktoren zu erwarten, anders als in der angeführten Untersuchung vonHANKE, LOHMÖLLER und MANDL (1978), die ihre Beurteilungen jeweils nach einem Schuljahr zu den Zeugnisterminen einholten, wenn sich die Urteiler aller Voraussicht nach in einem Zustand relativ großer subjektiver Urteilssicherheit befanden; als zusätzliche Vereinheitlichung der aktivierten IPTn der Lehrer mag dort noch die Vorgabe von Eigenschaftsbezeichnungen durch die Institution Schule gewirkt haben, die schon jahrelang als Kategorien der obligatorischen Schüler-Beurteilungsbogen gedient hatten. Wie die Autoren schon selbst schlußfolgern, deuten ihre hohen Strukturkongruenzen eine einheitliche Verwendung der von der Institution als beurteilungsrelevant vorgegebenen Begriffe durch die Lehrer an.

IV Methode

4.1. Theoretische Fundierung der Variablenerfassung

Welcher Art sollten nun die Informationen sein, mit deren Hilfe wir versuchen wollen, die korrelativen Urteilstendenzen in den Lehrerurteilen festzustellen ?

Stellen wir die Frage nach den Variablen, die die Beschaffenheit der Wahrnehmung beeinflussen, so stoßen wir auf das Modell sozialer Wahrnehmung von WARR u. KNAPPER (1968) (siehe Kapitel "Soziale Wahrnehmung"), auf das wir uns in den folgenden Ausführungen beziehen.

Hier werden Variablen in der Information, Variablen des Wahrnehmenden und variable Komponenten der Urteilsreaktion aufgeführt, die aber wechselseitig zueinander in Verbindung stehen und auf den Eingangsselektor und das Verarbeitungszentrum einwirken.
Da wir nicht den Wahrnehmungsprozeß selbst untersuchen wollen, sind Informationssammlungen, falls diese überhaupt in einer so komplexen Urteilssituation wie in dieser Feldstudie ausreichend geleistet werden könnten, in den meisten Variablenbereichen nicht geplant.
Wir wollen in dieser Arbeit ein Modell eines speziellen Systems von Inferenzregeln (WARR u. KNAPPER, 1968) erstellen, das wir im Verarbeitungszentrum lokalisieren. In dieser Arbeit sollen ferner keine Beziehungen zwischen Informationsvariablen, Variablen des Wahrnehmenden und dem Urteil festgestellt werden. Unsere Analyse bezieht sich nur auf gruppenstereotype Attribuierung von Eigenschaftszusammenhängen; die Datensammlung erstreckt sich also neben der reinen Stichprobenbeschreibung fast ausschließlich auf die attributiven Urteilsreaktionen.

4.2. Auswahl der Beurteiler-Stichprobe

Die Stichprobe bestand aus den 23 Lehrern aller ersten Klassen beider Gymnasien einer Kreisstadt in Baden-Württemberg mit ca. 50.000 Einwohnern. Für die Untersuchung der Verhältnisse in dieser Stadt repräsentiert die Stichprobe einen hohen Prozentsatz (74,19 %) der Population an Gymnasiallehrern der 1. Klasse, da nur die Lehrer von der Stichprobe ausgenommen wurden, die ausschließlich die Fächer Musik, Zeichnen, Turnen, Religion, gaben. Der Ausschluß erfolgte nicht aufgrund einer angenommenen geringeren Wertigkeit dieser Fächer, sondern weil einmal viele Vpn dieser Gruppe nicht zur Teilnahme zu motivieren waren, zum anderen, weil der Mittelwert der Stundenzahl, die sie in den Klassen verbrachten, (μ = 2,5/Woche) stark vom Mittelwert der Lehrer anderer Fächer (μ =5,15/Woche) unterschied, was mit großer Wahrscheinlichkeit zur Bildung einer Untergruppe von N = 8 geführt hätte.
Diese weitere Untergruppe mit geringerer Stundenzahl (= Kontaktzeit mit Schülern) wäre wegen des geringen Umfangs methodisch schlechter zu fassen gewesen und hätte mit großer Wahrscheinlichkeit die Werte der größeren Gruppe verzerrt. Zwar kann nicht angenommen werden, daß die Entwicklung ihrer Beurteilungen grundsätzlich anders verlaufen wäre; eine zeitliche Verzögerung, die aus selteneren Kontakten mit den Schülern resultiert, hätte aber die Ergebnisse der ganzen Studie verändert, wenn nicht gar den Abschluß gefährdet, da die Laufzeit der Studie durch die Rektoren auf 6 Monate begrenzt war.

4.2.1. Probleme der Stichprobe

Wie es in Feldstudien häufig der Fall ist, - besonders, wenn der Untersucher nicht mit der nötigen Amtsautorität ausgestattet ist, - hat der Untersucher keineswegs freie, nur durch seine Kenntnisse beschränkte Auswahl seines Vorgehens in Planung, Methodik und Durchführung der Studie. Vielmehr ist bei Feldstudien die Motivation der Vpn die dominierende Variable, die nur geringfügig beeinflusst werden kann, speziell in diesem Fall, in dem die Vpn bezüglich des institutionsfremden Untersuchers außerordentlich autonom waren. Die Variable Motivation oder deren Aufrechterhaltung waren an folgenden Brennpunkten zu berücksichtigen.

1.Problem der Auswahl der Vpn

Oft wird der Untersucher den Weg des geringsten Widerstandes gehen und nur solche Vpn auswählen, die sich spontan zur Teilnahme bereiterklären. Damit verbunden ist schon hier eine einseitige Selektion der Vpn. Zwang, sofern er ausgeübt werden kann, erzeugt sicher keine differenzierten Meßwertreihen. Außerdem vergrößert er nur das folgende Problem.

2. Problem der Mortalität der Stichprobe

Je kleiner der Stichprobenumfang, desto empfindlicher ist eine Untersuchung bezüglich dieses Problems, da oft die vorgesehene statistische Verrechnung der Daten unmöglich wird. Bei größeren Stichproben kann jedoch eine typendiagnostische Verwertung von "Aussteiger"-Daten wertvolle Aussagen zum Untersuchungsziel beitragen.

3. Problem des Aufwandes für die Versuchspersonen

Die beiden oben angesprochenen Problemkreise sind in ihrer Intensität größtenteils als Funktion des durch die Teilnahme bedingten Aufwandes für die Vpn zu sehen. Je geringer der Arbeitsaufwand ist, desto seltener treten die Probleme 1 und 2 auf.

Während in dieser Studie die ersten beiden Probleme durch persönliche Bekanntschaft des Untersuchers mit dem größten Teil der Vpn nicht auftraten oder wegen der sowieso nicht erreichbaren Repräsentativität irrelevant waren, forderte das Problem des Aufwandes einen Tribut in Form einer einschneidenden Reduktion der geplanten Methodik des Längsschnittes (Siehe Abschnitte über Auswahl der Beurteilten, Meßzeitpunkte und Meßinstrument).

4.2.2. Erfasste Variablen der Stichprobe

Erhoben wurden Angaben zu folgenden Merkmalen: Alter, Geschlecht, Anzahl der Dienstjahre, Hauptfach, Anzahl der Wochenstunden in einer Klasse.

Um eventuelle Affinitäten von Merkmalen der Urteilspersönlichkeit zur Verwendung bestimmter Skalen des Beurteilungsbogens festzustellen, wurde mit den Urteilern das FPI durchgeführt.
Es war noch die Durchführung des Fragebogens zur direktiven Einstellung (FDE) von BASTINE geplant, da diese Persönlichkeitsvariable als Teilkomponente autoritären Verhaltens im FPI nicht ausdrücklich gemessen wird. Die Skala 7 des FPI - "Dominanzstreben" - die den FDE hätte ersetzen können, wurde in der zweiten Auflage des Tests aufgrund weiterer Untersuchungen nur um die Bezeichnung "reaktive Aggressivität" erweitert, was sie in inhaltliche Nähe zur früheren Skala 2: "Aggressivität" (in der 2. Auflage "spontan-Aggressivität") bringt.

Das FPI (FAHRENBERG,SELG,1970 ist ein mehrdimensionaler Persönlichkeitstest, der itemmetrisch und faktorenanalytisch durchgearbeitet ist. Er besteht aus 212 items, die die folgenden 12 unterschiedlich langen Skalen bilden :

Skala 1:	Nervosität (psychosomatisch gestört	vs	psychosomatisch nicht gestört)
Skala 2:	Aggressivität (spontan aggressiv emotional unreif	vs	nicht aggressiv beherrscht)
Skala 3:	Depressivität (mißgestimmt selbstunsicher	vs	zufrieden selbstsicher)
Skala 4:	Erregbarkeit (reizbar, leicht frustriert	vs	ruhig, stumpf)
Skala 5:	Geselligkeit (gesellig lebhaft	vs	ungesellig zurückhaltend)
Skala 6:	Gelassenheit (selbstvertrauend gutgelaunt	vs	irritierbar zögernd)
Skala 7:	Dominanzstreben (reaktiv aggressiv sich durchsetzend	vs	nachgiebig gemäßigt)
Skala 8:	Gehemmtheit (gehemmt gespannt	vs	ungezwungen kontaktfähig)
Skala 9:	Offenheit (offen selbstkritisch	vs	verschlossen unkritisch)
Skala E:	Extroversion (extrovertiert	vs	introvertiert)
Skala N:	Emotionale Labilität (emotional labil	vs	emotional stabil)
Skala M:	Maskulinität (typ.männl.Selbstschilderung	vs	typ.weibl.Selbstschilderung)

Die Reliabilitäten der einzelnen Skalen sind insbesondere bei der hier verwendeten Gesamtform befriedigend. An Validitätsuntersuchungen interessiert besonders die Korrelation mit Selbsteinschätzungen, die z.T. relativ hohe Werte erbringen, wenn auch nicht immer mit den semantisch gleichen Skalen. Die Normierung erfolgte an einer für einen Test doch recht kleinen Stichprobe von N = 630).

In dieser Studie wurde die Gesamtform durchgeführt und nach der Handanweisung der 1. Auflage von 1970 ausgewertet.

4.3. Auswahl der Beurteilten-Stichprobe

Im allgemeinen werden bei der Untersuchung der IPTn die Beurteilten nur als Medium angesehen, anhand derer sich die IPTn manifestieren können, auf die sie projiziert werden können. Da aber ein Urteiler für verschiedene Gruppen von Beurteilten verschiedene Stereotypen hat, ist es doch sehr wahrscheinlich, daß für ebendiese Gruppen auch verschiedene IPTn vorliegen. (HASTORF et al.,1970)

Werden nun Stichproben aus Gruppen von Beurteilten zusammengesetzt, auf die der Urteiler jeweils andere IPTn anwendet, so ergibt die Datenanalyse ein Konglomerat aus allen spezifischen IPTn.
Aus den obengenannten Gründen (Siehe Abschnitt 4.2.) war es jedoch nicht möglich, dieser Fragestellung weiter nachzugehen und die Beurteilten-Stichproben zu homogenisieren oder nach vermuteten IPT-spezifischen Merkmalen auszulesen.
Es wurde vielmehr von den Vpn darauf gedrungen, die vorgesehenen Beurteilten-Stichproben (ganze Schulklassen) um 50 % zu reduzieren, um den zeitlichen Aufwand für die Beurteilungen gering zu halten.
Die Beurteilten-Stichprobe umfaßte schließlich je 20 zufällig ausgewählte Schüler und Schülerinnen aller ersten Klassen (= 140 Schüler), wobei die Schülerinnen 56,94 % der einzelnen Klassen ausmachten.
Ein Vergleich mit dem gleichen Meßwert aus den GRUND- UND STRUKTURDATEN für 1975 zeigt, daß dieser Prozentsatz weder für das Bundesland, in dem die Untersuchung stattfand, noch für das Bundesgebiet (47,9 %) typisch ist. Diese Erscheinung ist darauf zurückzuführen, daß eine der beiden Schulen in den Jahren vor der Untersuchung ein reines Mädchen-Gymnasium war, weswegen immer noch bevorzugt Mädchen auf diese Schule geschickt wurden.
Da eine Repräsentativität nicht erreichbar war, wurde die Stichprobe nach den Zahlenverhältnissen in der Grundgesamtheit aller Schüler der 1. Klassen beider Gymnasien am Ort organisiert, um wenigstens in den Aussagen über diesen Ort repräsentativ zu sein.

4.3.1. Erfasste Variablen der Stichprobe

Erhoben wurden Angaben zu Alter und Geschlecht der Beurteilten. Soziographische Daten wurden nicht erfasst, da eine Untersuchung damit zusammenhängender Fragestellungen nicht vorgesehen war; die beurteilenden Lehrer hatten offiziell ebenfalls keinen Zugang zu Daten dieser Art.

4.4. Methode zur Erfassung der impliziten Persönlichkeitstheorie durch Beurteilung von Stereotypen

Diese Methode hat auch HOFER (1969) angewandt, sodaß die Faktorenstruktur der 'subjektiven Erwartungswerte' (COHEN, 1969) als legitimes IPT-Modell angesehen werden kann.
Nach unseren Vorstellungen sollten die Lehrer der Stichprobe die Stereotypen "der gute Schüler", "der schlechte Schüler" und "der Schüler der ersten Gymnasialklasse" anhand eines Satzes von Rating - Skalen beurteilen. Die Pole der Rating - Skalen sollten mit Eigenschaftsbegriffen versehen sein, die von den Lehrern als relevant für die schulische Beurteilungssituation angesehen werden. Die aus diesem Plan resultierenden Verfahren werden in der Folge dargestellt.

4.4.1.1. Zwei typische Modelle in der 'person-perception' Forschung

Im allgemeinen werden in der umfangreichen Literatur zu Forschungszweigen der 'person perception', 'social perception','interpersonal perception', etc. im wesentlichen zwei typische Vorgehensweisen angegeben, um Urteile sichtbar und einer statistischen Erfassung und Verrechnung zugänglich zu machen.

1. Unstrukturierte Beurteilung

Vorteil: Die Beurteiler verwenden eigene Beurteilungskategorien, der Satz der verwendeten Beurteilungskategorien bildet am ehesten den persönlichkeitsspezifischen Gebrauch von Kategorien ab.

Nachteil: Das Feld der verwendeten Kategorien kann so umfangreich und so differenziert werden, daß nur noch eine individual-diagnostische Betrachtung möglich ist.
Falls der Untersucher jedoch zu Aussagen kommen möchte, die das rein individuell-deskriptive Niveau übersteigen, - was in den Sozialwissenschaften meistens angestrebt wird ist er nun doch gezwungen, zu interpretieren, zu abstrahieren, zu kategorisieren, - kurz, den Gewinn der anfangs unterlassenen Strukturierung durch subjektive, bias-intensive Manipulationen am Datenmaterial zu verspielen.

Noch gravierender wird der anfangs aufgeführte Nachteil, wenn gar noch mehrfache Beurteilungen mit zeitlichen Intervallen durchgeführt werden sollen. Hier kommt zu der bereits geschilderten Variationsbreite der Verwendung von Kategorien noch der mögliche Ausfall bisher verwendeter und das Verwenden neuer, bisher noch nicht benutzter Kategorien hinzu.
Weitere negative Einflüsse üben die unterschiedliche Verbalisierungsfertigkeit der Beurteiler und deren Motivation aus, an der Untersuchung teilzunehmen, mit dene die Menge und Güte der abgegebenen Urteile variiert.

2. Strukturierte Beurteilung

Vorteil: Durch die vorgegebenen Kategorien ist eine gute Vergleichsbasis und eine bessere Verbindung zu den dem Beurteilungsbogen zugrundeliegenden Hypothesen gegeben; die statistische Bearbeitung vereinfacht sich.
Der Beurteiler ist gezwungen, Urteile auf breiterer Basis abzugeben, als er es vielleicht gewohnt ist. Besonders, wenn weitere Beurteilungen in zeitlichen Abständen angekündigt werden, wird der Beurteiler im Kontakt mit den Beurteilten in den Intervallen evtl. gezielter auf die geforderten Kategorien hin beobachten, was durchaus einen Gewinn für seine Beurteilungspraxis darstellen kann, die sich dadurch von bisher praktizierter subjektiver Einseitigkeit lösen kann.

Nachteil: Gerade im Rahmen einer Untersuchung persönlichkeitsspezifischen Vorgehens bei Beurteilungen ist der zuletzt genannte Vorteil ein Nachteil. Der Beurteiler muß Kategorien verwenden, in denen er nicht beheimatet ist, in denen er demzufolge eine größere Urteilsunsicherheit aufweisen wird, denn gerade in diesen ungewohnten Kategorien ist ein besonders massiver Einfluß der bekannten Beurteilungsfehler wie dem "generosity error" (CRONBACH, 1960), dem "error of central tendency" (HOLLINGWORTH,1970, nach GUILFORD, 1954), dem "logical error" (NEWCOMB, 1931; GUILFORD, 1936) und dem "proximity error" (STOCKFORD u. BISSEL, 1949) festzustellen.

Die Ergebnisse aus Untersuchungen, die mit vorgegebenen Beurteilungsskalen arbeiten, können natürlich nichts anderes erbringen, als die vorgegebenen Kategorien vorsehen. Besonders, wenn die Kategorien vom Untersucher aufgrund einer Hypothese gezielt ausgesucht wurden, bewegen sich die Ergebnisse genau im deskriptiv zugelassenen Rahmen.

Betrachtet man einschlägige Arbeiten, die eine der beiden Methoden verwenden, so befriedigen die Ergebnisse beider Methoden in Hinsicht auf das in dieser Untersuchung angestrebte Ziel nicht.

Validität und Möglichkeit der Erkenntnis - sofern diese durch statistische Verarbeitung der Daten abgesichert werden soll - verhalten sich reziprok. Es müßte eine Hybrid-Methode gefunden werden, die beides ermöglicht.

Die Ergebnisse von HÖHN (1967), die auch mit der Methode der unstrukturierten Beurteilung bei ihren Urteilern nur eine geringe Variabilität der verwendeten Kategorien feststellte, lassen es möglich erscheinen, durch einen begrenzten Satz vorgegebener Kategorien doch die Variationsbreite individueller Beurteilungsschemata wenigstens einer homogenen Urteilergruppe in Bezug auf eine homogene Beurteiltengruppe zu erfassen.
Besonders genährt wird diese Annahme noch durch die Tatsache, daß sich Urteiler und Beurteilte ausschließlich in fixierten Rollen in einem institutionell fixierten Raum begegnen.

Gefestigt wurde der Entschluß zur Vorgabe von Rating-Skalen einmal nach den Ergebnissen von KIEFERLE u. SECHREST (1961), nach denen sich das Urteilsverhalten bei vorgegebenen Skalen nur in geringem Maß von dem unterschied, was bei freier Wahl der Kategorien praktiziert wurde, zum anderen durch BOGEN (1927), der bei seinen Lehrern eine Bevorzugung von vorgegebenen Beurteilungsskalen feststellte. Allerdings sollten diese vorgegebenen Skalen keine aus Wörterbüchern zusammengesuchten Kunstprodukte sein, deren Relevanz für die Vpn der Stichprobe nur eine angenommene ist; die Vpn sollten vielmehr ihre individuellen Vorstellungen in die Konstruktion der Skalen eingehen lassen können, ein Vorgehen, das sich durch SLOCOMBE (1938) bestätigt fand.

4.4.1.2. Erhebung relevanter Eigenschaftsbegriffe

Die Auswahl der Eigenschaftsbegriffe soll also in unserer Arbeit nicht vom Untersucher vorgenommen werden, der oftmals sein eigenes Sprachgefühl als Maß für die Bedeutung der Wörter verwendet. Wir stellten daher eine Reihe von Fragen an die Lehrer, die einen Satz von Eigenschaften erbringen sollten, die die Lehrer selbst als relevant für die Beurteilungssituation erachteten.

Frage 1: "Nach welchen Merkmalen beurteilen Sie die Persönlichkeit Ihrer Schüler? Bitte benutzen Sie Eigenschaftswörter."

Diese Frage sollte Eigenschaften erbringen, die die Lehrer als relevant für die Beurteilung von Schülern ansahen. Durch den Zwang, Adjektive zu benützen, wird die Verwendung anderer Urteilskategorien (etwa in Substantiv- oder Satzform) unmöglich, was von HUBER/MANDL (1979) kritisiert wird und die Validität solcher Untersuchungen natürlich empfindlich senkt. Andererseits muß keine Manipulation mit dem Rohmaterial durchgeführt werden, um eine Vergleichsbarkeit zu erzeugen.

Jedes genannte Merkmal wurde einmal in der Reihenfolge der Nennung notiert (spontane Rangreihe) und zusätzlich auf eine Karteikarte geschrieben. (Ergebnisse s. Tab. 1; Vergleich der erhaltenen Eigenschaftsbegriffe mit den von anderen Autoren ausgesuchten: s.Tab.2).

Frage 2: "Bringen Sie diese Merkmale bitte in eine Rangreihe nach Ihrer Wichtigkeit für Sie; welches Merkmal ist für Sie selbst, Ihrer Meinung nach, das wichtigste, das zweitwichtigste,...".

Die unter 1) hergestellten Merkmalskarten waren inzwischen unauffällig gemischt worden und die Vp legte nun eine Rangreihe, die schriftlich fixiert wurde. Das Ergebnis soll als "reflektierte Rangreihe" bezeichnet

Tabelle 1

Frage 1: Nennungshäufigkeiten der Beurteilungskategorien

Kategorie	f
intelligent	23
begabt	23
gute Mitarbeit	23
fleißig	23
ordentlich	17
sprachlich begabt	15
konzentriert	14
selbständig	14
eigenes Mitdenken	14
ehrgeizig	10
gutes häusl. Milieu	10
stört nicht Unterricht	9
zuverlässig	9
positiv eingestellt zu Schule und Lernen	9
ausdauernd	9
offener Charakter	7
aufgeweckt	7
kontaktfreudig	5
vital	4
höflich	4
sensibel	4
überfordert	4
kreativ	3
paßt sich Klasse an	3
guter Lehrerkontakt	3
ehrlich	2
akzeleriert	2

$\leq$ 1: sportlich, musisch begabt, klares Denken, verspielt, von zu Hause gezwungen, langsam, begreift Lehrstoff, anständig, saubere Heftführung, pünktlich.

(1) HÖHN		(2) HOFER		(3) VORLIEGENDE UNTERSUCHUNG	
begabt	- unbegabt	begabt	- unbegabt	begabt eher math. begabt	- unbegabt - eher sprachl. begabt
aufgeschlossen	- geistig eng				
interessiert	- uninteressiert	interessiert	- uninteressiert	positiv eingest. zu Schule u. Lernen	- negativ eingest. zu Schule u. Lernen
geschickt	- ungeschickt				
ruhig	- unruhig	ruhig	- unruhig		
gesammelt	- zerstreut	aufmerksam	- unaufmerksam		
		konzentriert	- unkonzentriert	konzentriert	- unkonzentriert
schnell	- langsam				
leise	- laut				
sorgfältig	- nachlässig	ordentlich	- unordentlich	ordentlich	- schlampig
		pflichtbewußt	- pflichtvergessen	zuverlässig	- unzuverlässig
fleißig	- faul	fleißig	- faul	fleißig	- faul
zart	- grob	sensibel	- robust	sensibel	- seelisch robust
gut	- böse				
froh	- traurig				
offen	- verschlossen	offen	- verschlossen	offener Charakter	- verschlossener Charakter
ehrlich	- unehrlich				
brav	- ungezogen	folgsam	- unfolgsam		
bescheiden	- stolz	bescheiden	- geltungsbedürftig		
		schüchtern	- dreist		
mutig	- feig				
kameradschaftl.	- unkameradschaftl.				
gelöst	- gespannt	ausgeglichen	- launisch		

Fortsetzung Tab. 2

(1) HÖHN		(2) HOFER		(3) VORLIEGENDE UNTERSUCHUNG	
natürlich	- eitel				
gepflegt	- verwahrlost				
groß	- klein			gutes häusl. Milieu	- schlechtes häusl.Milieu
stark	- schwach			akzeleriert	- retardiert
gesund	- krank			vital	- schwächlich
hübsch	- häßlich			(hübsch	- nicht hübsch)
		ehrgeizig	- gleichgültig	ehrgeizig	- nicht ehrgeizig
		einfallsreich	- einfallslos		
		führend	- zurückhaltend		
		gesellig	- ungesellig	kontaktfreudig	- kontaktarm
		höflich	- frech	höflich	- frech
		intelligent	- dumm	intelligent	- unintelligent
		unkompliziert	- kompliziert		
		selbstsicher	- selbstunsicher		
		(sympathisch	- unsympathisch)	sympathisch	- nicht sympathisch
				unterfordert	- überfordert
				paßt sich Klassengemeinschaft an	- paßt sich Klassengemeinschaft nicht an
				gute Mitarbeit	- schlechte Mitarbeit
				stört nie	- stört häufig
				eigenes Mitdenken	- kein eigenes Mitdenken
				selbständig	- unselbständig
				guter Lehrerkontakt	- kaum Lehrerkontakt

werden, da die Vpn im Gegensatz zu Frage 1 gezwungen waren, sich über die Vergabe der Rangplätze Gedanken zu machen, zu gewichten, ihr Urteilsverhalten zu reflektieren. Diese Rangreihe soll dem Vergleich mit'zentralen Eigenschaften' dienen. (Ergebnis: s.Tab.3).

Sehen wir uns Tabelle 2 an, so stellen wir fest:

Die rein sprachliche Gegenüberstellung zeigt die unterschiedlichen Auffassungen der Autoren über die Relevanz der zur Beurteilung bereitgestellten Kategorien, einmal an der Zahl der "überzähligen" Kategorien, für die es zwischen den Untersuchungen keine Entsprechungen gibt, zum anderen an der Differenz des Gegenpols bei gleichlautenden Kategorien (z.B. HÖHN, 1967,: "bescheiden-stolz", HOFER, 1969,: "bescheiden-geltungsbedürftig"). Außerdem kann man sich des Gefühls eines gewissen künstlichen Sprachgebrauchs nicht erwehren; Worte wie "selbstunsicher" (HOFER, 1969) als konstruierter Gegenpol zu "selbstsicher" spiegeln zwar die Besorgnis des Autors wieder, nur ja nicht durch unkontrollierte Implikationen bei der Wahl eines gebräuchlicheren Wortes die Eindimensionalität der Skala zu gefährden, führen aber letzten Endes doch dazu, wenn dieser konstruierte Gegenpol von den Vpn nicht als Gegenpol des eigenen Sprachgebrauchs angenommen wird.

Um diesen Erscheinungen zu entgehen, die auch die Kritik von GIGERENZER u. STRUBE (1978, S.79) hervorgerufen haben, wird zur Herstellung von Skalen, die - von den Vpn her gesehen - subjektive Eindimensionalität aufweisen, die folgende Frage an die Lehrer gestellt.

Frage 3:" Könnten Sie bitte das Gegenteil der von Ihnen genannten Eigenschaft .... nennen?"

Die Ergebnisse (s.Tab.3) zeigten eine weitestgehende Übereinstimmung der Nennungen, was auf eine weitgehende Normierung des Sprachgebrauches in der Stichprobe schließen lässt. Da wir aber annehmen, daß der gruppenspezifische Sprachgebrauch im Verlauf der Sozialisation gelernt wurde, überrascht dieses Ergebnis nicht. Nach KELLY (1955), der mit Nennungen der Gegenteile vorgegebener Eigenschaften die Konnotationen von Begriffen zu erhalten versuchte, könnten diese hohen Übereinstimmungen der Gegenteilsnennungen nach diesem etwas groben Maß eine weitgehende Ähnlichkeit der affektiven Wertungen der vorgegebenen Begriffe andeuten.

Die nächsten Fragen wurden nach der Zusammenstellung der erfragten Eigenschaftsbegriffe in Rating - Skalen und dieser zu Listen (siehe nächster Abschnitt) gestellt und dienten dem Ziel der Erhebung der Stereotypen, aus denen die Faktorenstruktur errechnet werden sollte.

Frage 4:"Im Folgenden werden Sie gebeten, Ihr Urteil über "den" Schüler der 1. Klasse im Gymnasium abzugeben. Natürlich gibt es "den Schüler" der 1. Klasse nur als theoretisches Konstrukt, als Mittelwert; es soll hier aber nach der Essenz Ihrer jahrelangen Erfahrung mit Schülern der 1. Klasse gefragt werden,wie sind die meisten von ihnen? Welches Bild haben Sie also in diesem Sinn von "dem" Schüler der 1. Klasse, der Ihnen heute im Schulalltag gegenübersteht?"

Tabelle 3

Frage 2:	Frage 3:		
Rangordnung	genanntes Gegenteil	Übereinstimmung in %	absolut
1. gute Mitarbeit	schlechte Mitarbeit	100	23
2. intelligent	nicht intelligent	91,1	21
	dumm	8,9	2
3. fleißig	faul	95,7	22
	bequem	4,3	1
4. begabt	unbegabt	95,7	22
	dumm	4,3	1
5. ordentlich	schlampig	86,9	20
	unordentlich	13,1	3
6. konzentriert	unkonzentriert	91,1	21
	abschweifend	8,9	2
7. stört nie den Unterricht	stört oft den Unterricht	78,1	18
	stört immer	21,9	5
8. zuverlässig	unzuverlässig	100	23
9. sprachlich begabt	mathematisch begabt	82,5	19
	sprachlich unbegabt	17,5	4
10. eigenes Mitdenken	kein eigenes Mitdenken	78,1	18
	passiv	17,6	4
	lethargisch	4,3	1
11. ausdauernd	nicht ausdauernd	82,5	19
	erlahmt rasch	13,2	3
	gibt schnell auf	4,3	1
12. positiv eingestellt zu Schule und Lernen	negativ eingestellt zu Schule und Lernen	91,1	21
	haßt Schule u. Lernen	8,9	2
13. selbständig	unselbständig	73,8	17
	muß angeleitet werden	26,2	6
14. kontaktfreudig	kontaktarm	73,8	17
	scheut Kontakte	17,5	4
	schließt kaum Kontakte	8,7	2
15. ehrgeizig	nicht ehrgeizig	100	23
16. höflich	frech	86,9	20
	unhöflich	13,1	3
17. aufgeweckt	geistig träge	69,6	16
	verschlafen	30,4	7
18. kreativ	nicht kreativ	73,8	17
	nicht originell	17,5	4
	hat keine neuen Ideen	8,7	2
19. gutes häusl. Milieu	schlechtes häusl.Milieu	87,1	20
	wird zu Hause nicht unterstützt	8,6	2
	verwahrlost	4,3	1

Frage 3: Rangordnung	Frage 3: genanntes Gegenteil	Übereinstimmung in %	absolut
20. offener Charakter	verschlossener Charakter	74,0	17
	verschlossen	21,7	5
	mürrisch	4,3	1
21. vital	schwächlich	69,6	16
	sensibel (!)	30,4	7
22. sensibel	seelisch robust	78,3	18
	robust	13,0	3
	vital (!)	8,7	2
23. unterfordert	überfordert	95,7	22
	gefordert	4,3	1
24. paßt sich Klasse an	paßt sich Klasse nicht an	87,1	20
	hat einen eigenen Kopf	8,6	2
	schließt sich aus	4,3	1
25. guter Lehrerkontakt	schlechter Lehrerkontakt	100	23
26. akzeleriert	retardiert	91,3	21
	körperlich zurückgeblieben	8,7	2
27. ehrlich	unehrlich	91,3	21
	verlogen	8,7	2

Tabelle 3: Rangreihe der Eigenschaftsnennungen (Frage 2) sowie Nennungen der sprachlichen Gegenteile (Frage 3) und ihre absolute und relative Übereinstimmung in der vorliegenden Stichprobe.

Die Fragen 5) und 6) bezogen sich in gleicher Formulierung auf "den guten" und "den schlechten" Schüler der ersten Gymnasialklasse.

4.4.1.3. Herstellung der Rating-Skalen

4.4.1.3.1. Kennzeichnung der Pole

Aus der Primärliste der auf Frage 1 erhaltenen Eigenschaftsbegriffe wurden alle, die nur ein einziges Mal genannt wurden, ausgeschieden. Diese Kategorien befanden sich ausnahmslos am Ende der nach Frage 2 gebildeten Rangreihen, stellten also wohl kaum subjektiv zentrale Urteilskategorien der Beurteiler dar.
Zwar wird durch dieses Ausscheiden selten genannter Eigenschaften die Möglichkeit geringer, Variabilität abzubilden, was sich validitätssenkend auf unser IPT-Modell auswirkt; HOFER (1969, S.86) fand zwischen der Anzahl der verwendeten Eigenschaften und der Zahl der extrahierten Faktoren eine Korrelation von .90 (!). Einerseits ist zu überlegen, daß wir zwar durch diesen Eingriff die Anzahl der resultierenden Faktoren wahrscheinlich senken werden, daß wir andererseits aber die Rating-Listen von 10 Skalen säubern, die nur für einen Bruchteil der Urteiler relevant ist - und für diese auch nur an sehr nachgeordneter Stelle.

Die resultierenden 27 Skalen wurden vom Untersucher noch um zwei Skalen erweitert; bereits in Frage 3 waren die Eigenschaftsbegriffe "sympathisch" und "hübsch" unter die von den Lehrern genannten gemischt worden, sodaß auch hier die Definition der Gegenteile von den Lehrern vorgenommen wurde.

Reizwort	Gegenteil	Übereinstimmung in%	absolut
sympathisch	unsympathisch	100	23
hübsch	nicht hübsch	52,2	12
	häßlich	39,1	9
	unscheinbar	8,7	2

Gerade bei der Vorgabe des Eigenschaftsbegriffs "hübsch" resultierten die größten Differenzen in den Gegenteilsnennungen, was uns für das Vorhaben dieses Eingriffs hoffen ließ. Diese "sachfremden" Skalen sollten nämlich in die Liste aufgenommen werden zur Feststellung globaler Wertungen in den Urteilen der Lehrer und speziell auch zur Beantwortung der Frage, welche der "sachbezogenen" Eigenschaften emotional geladen sind; man könnte annehmen, daß sich diese Eigenschaftsbezeichnungen als "zentrale Eigenschaften" im Sinne von ASCH (1946) etablieren; andererseits wurden sie von den Lehrern als nicht relevant erachtet, bzw. wurden sie als sozial nicht anerkanntes Beurteilungskriterium im Schulbereich einem Fremden gegenüber nicht genannt.

4.4.1.3.2. Skalierung

Einer weitverbreiteten Konvention bei der Anwendung von Rating-Skalen in der Forschung zur 'social-perception' folgend, nehmen wir unberechtigterweise Intervallskalen-Niveau an. Eine Diskussion neuer Ansätze wie dem von HUBER u. MANDL (1979), die diese und die Vorgabe von Eigenschaften betreffende Implikationen vermeiden, soll im Kapitel "Diskussion der Ergebnisse" stattfinden. Die vorliegende Arbeit ist als weitgehende Replikation zur Arbeit von HOFER (1969) und zu deren kritischer Beleuchtung von den Ergebnissen der aktuellen Schülerbeurteilungen her, soweit das von der Methodik her möglich war, zu betrachten.

Die aus den Fragen 1 und 3 resultierenden Eigenschaftsbegriffe wurden als Pole 7-stufiger Skalen eingesetzt. Die Stufen der Skalen wurden mit 3 2 1 0 1 2 3 bezeichnet. Von einem Vorzeichen (+,-) wurde abgesehen, um keine Wertung einzelner Pole zu suggerieren.
Bei der nun folgenden Beurteilung der drei Stereotypen wurde für jeden Lehrer eine andere Rating-Liste entworfen, in der sowohl die Reihenfolge als auch die Stellung der Adjektive links oder rechts neben der Zahlenreihe nach Zufall verändert worden war, um seriale Effekte zu verhindern.

4.4.1.3.3. Auswahl der Stereotypen

Zur Datensammlung ließen wir die drei Stereotypen "der gute Schüler", "der Schüler der 1. Gymnasialklasse" und "der schlechte Schüler" in zufälliger Reihenfolge beurteilen.
Der Stereotyp "schlechter Schüler" und "guter Schüler" sind symmetrisch um "den Schüler" angeordnet, um keine einseitig verzerrten Mittelwerte auf einzelnen Skalen zu provozieren und damit den später geplanten Vergleich der Faktorenstruktur mit der aus aktuellen Schülerbeurteilungen nicht zu belasten.

HOFER (1969, S.28) verzieht aller Wahrscheinlichkeit nach die Mittelwerte einzelner Skalen durch die Beurteilung eines vierten Stereotyps ("idealer Schüler") zur positiven Seite hin. Das wird sich zwar auf die Interkorrelationen der Eigenschaftsbegriffe und somit auf die Faktorenstruktur nicht so gravierend auswirken, wohl aber auf eine "zentrale Eigenschaft", (mit WISHNER, 1960, definiert als die Skala, die die meiste Varianz bindet, statistisch ausgedrückt durch die Skala mit der höchsten Kommunalität). Wir könnten als Indiz für das Zutreffen dieser Überlegungen wenigstens mit einer besonderen Zentralität einer Eigenschaftsskala rechnen, die in einer Affinität steht zu einem eher positiven durchschnittlichen Schülerbild (Stereotyp). Tatsächlich schiebt sich beim Vergleich der Faktorenstruktur der Stereotypen mit der Faktorenstruktur der neutraleren Ähnlichkeitsschätzungen von Eigenschaften die Skala "sympathisch - unsympathisch" mit einer hohen Kommunalität in den Vordergrund der Struktur der Stereotypen, während sie in der Struktur der Ähnlichkeitsschätzungen eine geringe Bedeutung (=Kommunalität) für die Lehrerurteile hat. (HOFER, 1969, S.77).
Diese Skala war eingeführt worden, um den emotionalen Bezug der Lehrer zu den einzelnen Eigenschaften festzustellen. Wir können also schlussfolgern, daß der asymmetrische Aufbau der Stereotypen-Beurteilungen einen Einfluß auf den Vergleich der Faktorenstrukturen hatte, wenn auch die Auswirkungen nicht so gravierend waren, daß die Eigenschaftsskala mit zentraler werdender Bedeutung soviel Varianz bindet, um einen eigenen Faktor zu determinieren.

4.4.1.3.4. Statistische Bearbeitung der Daten

Aus den Beurteilungen wurde analog zum Vorgehen in den Vergleichsarbeiten für jeden Lehrer auf jeder Skala der Mittelwert der drei Beurteilungen errechnet und als Meßwert verwendet; die Eigenschaftsskalen wurden interkorreliert und die Eigenschaften x Eigenschaften-Matrix einer Faktorenanalyse unterzogen.

4.5. Methode zur Erfassung der Faktorenstruktur aktueller Schülerbeurteilungen: Der Längsschnitt

4.5.1. Probleme der Datensammlung in Längsschnittuntersuchungen

4.5.1.1. Design-Modelle

HELMREICH (1977, S.11 f) stellt folgende Stichproben-Designs dar:

1. Trenduntersuchungen

Zur Verfolgung von Trends wird zu jedem Erhebungszeitpunkt eine neugebildete Personenstichprobe aus der Population gezogen, die mit der vorhergehenden nicht identisch ist. Dieser Typ wird häufig bei Meinungsumfragen verwendet, wobei sich z.B. bei wiederholten Repräsentativumfragen sogar die Population ändern kann.

Aussagemöglichkeit: Feststellung durchschnittlicher Veränderung. Keine empirische Erklärung für diese Veränderung möglich.

2. Kohorten-Studie

Hier sind die Merkmale der Population, aus der die Personenstichproben gezogen werden, genauer beschrieben als bei der Trenduntersuchung. Die Stichproben werden zu allen Durchführungszeitpunkten aus einer Population gezogen, die in ihren Merkmalen den Populationen früherer Stichproben gleicht.

Aussagemöglichkeit: Feststellung durchschnittlicher Veränderung in genau beschriebenen Gruppen.

3. Panel-Untersuchungen

Zu jedem Meß-Zeitpunkt wird eine identische Personen-Stichprobe (Panel) aus der gleichen Population gezogen, wobei identische Merkmale gemessen werden.

Aussagemöglichkeit: Zusammenhänge und Unterschiede der Daten zwischen den Zeitpunkten können nicht nur als Trend festgestellt werden, es können vielmehr personspezifische Entwicklungen verfolgt werden.

4. Simulierte Längsschnittuntersuchung aus einem Querschnitt-Design

Es erfolgt nur eine Erhebung zu einem Zeitpunkt; Trend-Untersuchungen lassen sich daraus simulieren, indem Daten vergleichbarer Stichproben verschiedener Alters- und Ausbildungsstufen längsschnittanalog verglichen werden.

Aussagemöglichkeit: Unterschiede zwischen Probandengruppen werden als Veränderung interpretiert.

5. Follow-up-Untersuchungen

Während bei Panel-Untersuchungen die Datensammlung in der Gegenwart beginnt und zukünftig wiederholt wird, greifen Follow-up-Untersuchungen auf zeitlich vor der Untersuchung entstandene Daten zurück.

Unsere Untersuchung ist demnach als Panel-Untersuchung zu betrachten.

4.5.1.2. Probleme der Stichprobe

Bereits im Kapitel 3.2. erfolgte ein Vorgriff auf diese Problematik. In vielen Untersuchungen (so auch bei HOFER, 1969) ist die Motivation der Lehrer zur Teilnahme ein folgenreiches Auslesekriterium, das allerdings nicht nur in Längsschnitt- und/oder Feldstudien besteht. Durch dieses Auslesekriterium entsteht eine einseitig selegierte, unrepräsentative Stichprobe, die die Generalisierbarkeit der Ergebnisse empfindlich senkt.
Spezifisch für Längsschnittuntersuchungen - und das gilt in verstärktem Maße für Feldstudien mit diesem Design - ist jedoch die Problematik der Aufrechterhaltung der Motivation über die Zeit hinweg und bei längeren Untersuchungen das Wiederauffinden der Probanden. Dieses letzte Problem betrifft eher eine Untersuchungs-Situation, in der kein ständiger Kontakt zwischen Untersucher und Probanden aufrechterhalten wird, mithin also nicht die vorliegende Arbeit.
Die Rekrutierung der Lehrer als Probanden und die Verhinderung der Stichproben-Mortalität aus Gründen geringer oder im Verlauf der Untersuchung abnehmender Motivation war problematisch. Die Motivation erwies sich als inhalts- und aufwandsabhängig. Kurz vor Beginn der Untersuchung war wieder einmal eine der periodisch zu Schuljahresende auftretenden Presse-Kampagnen bezüglich der Subjektivität der Leistungsbeurteilungen unter Bezug auf psychologische Untersuchungen (!) durchgeführt worden. Der Autor wurde sogar verdächtigt, in dieser Hinsicht Material zu sammeln. Erst nach dem Beweis, daß keinerlei Überprüfung der Notengebung erfolgte, konnten die vorgesehenen Lehrer dazu bewogen werden, die Rating - Listen auszufüllen.

Sehr viel gravierender auf die Motivation wirkte sich jedoch der mit der Untersuchung für den einzelnen Lehrer verbundene Aufwand aus, was weitgehende Konzessionen und ein Abweichen vom ursprünglichen Versuchsplan nötig machte, um die Untersuchung überhaupt durchführen zu können. Diese Konzessionen waren:

- nicht alle Schüler und Schülerinnen der Klasse wurden beurteilt, sondern nur 20 pro Klasse
- die Rating-Listen wurden nicht wöchentlich ausgefüllt, sondern nur alle zwei Wochen
- über die Schüler-Rating-Listen hinaus dürfen keine weiteren, mit Zeitaufwand für die Lehrer verbundenen, Datenerhebungen durchgeführt werden. Dem fiel eine geplante Sammlung von Daten über die Kontext-Wahrnehmung (Einschätzung der ganzen Klasse auf den Rating-Listen) sowie eine lose Sammlung von Informationen der Lehrer über jeden beurteilten Schüler zum Opfer.
- die Gesamtdauer der Untersuchung wurde auf ein Schulhalbjahr beschränkt, mit der Möglichkeit einer einmaligen Nachuntersuchung.

Diese Einschränkungen ließen einige Fragestellungen, die die Untersuchung etwas interessanter hätten machen können, nicht zu, z.B. das Vorhaben, die Beziehung zwischen dem sich allmählich im Verlauf des Kontakts aufbauenden Informationshintergrund des Lehrers bezüglich eines Schülers und dem "Bild", das er sich von diesem Schüler machte. zu verfolgen. Der Untersucher war aber mit keinerlei Amtsautorität versehen, sodaß an ein Durchsetzen einzelner Notwendigkeiten nicht zu denken war.

4.5.1.3. Folgen wiederholter Datenerhebung

Es kann vermutet werden, daß eine häufige Konfrontation der Lehrer mit den Rating-Listen eine Beeinflussung der Stichprobe darstellt, die die Generalisierung der Ergebnisse nicht zulässt.
Diesen Anspruch haben wir jedoch schon aus theoretischen Überlegungen aufgegeben. Das Auftreten eines solchen "Testungseffekts" könnte in einer Sensibilisierung der Lehrer für bestimmte Inhalte der Rating-Skalen begründet sein. Damit soll nicht eine Rückerinnerung an frühere Markierungen einzelner Schüler-Ratings gemeint sein - dazu sind doch die Intervalle zwischen den Beurteilungszeitpunkten zu groß -, sondern daß durch die Anzahl der in den Rating-Listen angebotenen Kategorien eine Sensibilisierung für darin angesprochene Sachverhalte erfolgt, die - bei Vertreten des realistischen Interpretationsmodus der IPT - eine Veränderung der Faktorenstruktur oder -inhalte bewirken könnte. Hier sind wir wieder bei dem bereits diskutierten Thema der Relevanz, mit der einzelne Urteilskategorien dem einzelnen Urteiler erscheinen. Wir haben zwar in unserer Untersuchung eine Relevanz aller Urteilskategorien (Eigenschaftsbegriffe) durch die Erhebung von den Lehrern selbst aufzuweisen, eine Relevanz betrifft aber nur die gesamte Gruppe der Lehrer, nicht einzelner Lehrer. Dem einzelnen Lehrer mögen zwar alle Eigenschaftsbegriffe als relevant für eine Schülerbeurteilung erscheinen, er selbst benutzt aber einen großen Teil von ihnen gewöhnlich nicht bei der impliziten Beurteilung seiner Schüler - falls er sich überhaupt auf dem Niveau von Eigenschaftsbegriffen bewegt; wie schon mehrfach andiskutiert wurde, bieten inhaltsanalytische Verfahren nur eine Teillösung des Validitätsproblems.

Um wenigstens den "Testungseffekt" abschätzen zu können, der durch die Häufigkeit der Erhebungen bedingt ist, könnte man Kontrollgruppen einführen, die mit größeren zeitlichen Abständen untersucht werden und Querschnittsvergleiche zwischen Versuchsgruppe und Kontrollgruppe anstellen, was jedoch ebenfalls problematisch ist, wenn keine vergleichbare Probandengruppe zusammengestellt werden kann.
Auf die Problematik der Meßmethoden in Längsschnittuntersuchungen wird in einem späteren Abschnitt eingegangen.

4.5.2. Die Urteiler-Stichprobe (Lehrer)

Da aus theoretischen Überlegungen nur eine Stichprobe von Lehrern vorgesehen war (siehe Kapitel 4.3., Diskussion der Frage 2), die sowohl die Stereotypen-Beurteilung als auch die Beurteilung realer Schüler vornehmen sollte, wurde die Auswahl der Stichprobe bereits beschrieben (siehe Abschnitt 4.2.). Da wir die Zahl möglicher Kontakte zwischen Lehrern und beurteilten Schülern im Schulbereich wenigstens geringfügig normieren wollten, waren also 8 Lehrer der Nebenfächer mit durchschnittlich 2,5 Wochenstunden in der Klasse aus der Stichprobe genommen worden; diese Maßnahme hat sicher in einer so komplexen Untersuchungssituation keinen Einfluß auf eventuelle Wahrnehmungshäufigkeiten; sie sollte nur einer gewissen Homogenisierung der Stichprobe dienen.

4.5.2.2. Die Beurteilten-Stichprobe (Schüler)

Die Schüler-Stichprobe umfasste je 20 nach Zufall ausgewählte Schüler und Schülerinnen aller sieben ersten Klassen des Ortes (=140 Schüler), wobei die Schülerinnen 45% der Stichprobe stellten.
Ein Vergleich mit den GRUND-UND STRUKTURDATEN (1976, S.28) zeigte keine signifikante Abweichung der ganzen Stichprobe vom Bundesdurchschnitt, während in einzelnen Klassen ein Überwiegen der Mädchen festzustellen war.
Diese Tatsache scheint sich aber nicht verzerrend auf die Ergebnisse auszuwirken, da HANKE, LOHMÖLLER und MANDL (1978, S.82 ff) keinen Einfluß des Geschlechts der Beurteilten auf die Sichtweise des Urteilers feststellten.
Soziographische Daten der Schüler wurden nicht erhoben, da keine nach diesem Kriterium differenzierte Datenbearbeitung geplant war. Während allerdings HUBER,LOHMÖLLER und MANDL (1978, S.85 ff) keine mit der Schichtzugehörigkeit der Beurteilten korrespondierende spezifische IPT feststellten, kommt SCHWARZER (1976) aufgrund der Ergebnisse ihrer Untersuchung zu dem Schluß einer Relevanz des sozioökonomischen Status der Schüler für die soziale Wahrnehmung'des Lehrers'(a.a.O., S.79).

Um einen O-Punkt gespeicherter Information über die Schüler der Stichprobe anzustreben, wurden die Lehrer über die Bekanntschaft mit einzelnen Schülern befragt. Diese Schüler wurden aus der Beurteilten-Stichprobe genommen und durch unbekannte ersetzt.

4.5.3. Das Meßinstrument

4.5.3.1. Die Rating-Liste

Für die Beurteilung realer Schüler wurde die gleiche Liste mit Rating-Skalen verwendet wie für die Beurteilung der Stereotypen (s.Abb. 3). Von jedem Meßzeitpunkt zum anderen wurde die Reihenfolge der Eigenschaftsbegriffe und die Polung nach Zufall verändert, um seriale Effekte auszugleichen. Zusätzlich wurde noch die im nächsten Abschnitt angesprochene Besonderheit eingefügt.

..... Lehrer-Nr.:.....
Wochen-Nr.:.....

hübsch	3 2 1 0 1 2 3	nicht hübsch
vital	3 2 1 0 1 2 3	schwächlich
ordentlich	3 2 1 0 1 2 3	schlampig
unkonzentriert	3 2 1 0 1 2 3	konzentriert
kreativ	3 2 1 0 1 2 3	nicht kreativ
guter Kontakt zum Lehrer	3 2 1 0 1 2 3	schlechter Kontakt
positiv eingestellt zu Schule und Lernen	3 2 1 0 1 2 3	negativ eingestellt
paßt sich der Klassen-gemeinschaft nicht an	3 2 1 0 1 2 3	paßt sich an
unbegabt	3 2 1 0 1 2 3	begabt
kontaktfreudig	3 2 1 0 1 2 3	kontaktarm
nicht ausdauernd	3 2 1 0 1 2 3	ausdauernd
retardiert	3 2 1 0 1 2 3	akzeleriert
gute Mitarbeit	3 2 1 0 1 2 3	schlechte Mitarbeit
seelisch robust	3 2 1 0 1 2 3	sensibel
stört häufig	3 2 1 0 1 2 3	stört selten
frech	3 2 1 0 1 2 3	höflich
faul	3 2 1 0 1 2 3	fleißig
eigenes Mitdenken	3 2 1 0 1 2 3	kein eigenes Mitdenken
schlechtes häusl. Milieu	3 2 1 0 1 2 3	gutes häusl. Milieu
unselbständig	3 2 1 0 1 2 3	selbständig
ehrlich	3 2 1 0 1 2 3	unehrlich
unsympathisch	3 2 1 0 1 2 3	sympathisch
aufgeweckt	3 2 1 0 1 2 3	geistig träge
verschlossener Charakter	3 2 1 0 1 2 3	offener Charakter
überfordert	3 2 1 0 1 2 3	unterfordert
nicht ehrgeizig	3 2 1 0 1 2 3	ehrgeizig
eher sprachl.begabt	3 2 1 0 1 2 3	eher mathem. begabt
zuverlässig	3 2 1 0 1 2 3	unzuverlässig
intelligent	3 2 1 0 1 2 3	nicht intelligent
.....	3 2 1 0 1 2 3	.....

Wie sicher sind Sie sich Ihres Urteils?

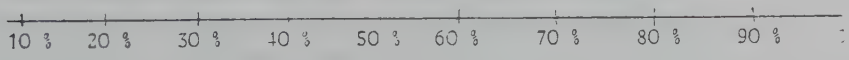


Abb.3: Muster einer Rating-Liste

4.5.3.2. Skala der "subjektiven Sicherheit"

Mit der besonderen zeitlichen Anlage des Längsschnitts, nach der ein allmählich steigender Bekanntheitsgrad zwischen Urteilern und Beurteilten und deren Wirkung auf die Faktorenstruktur aktueller Beurteilungen verfolgt werden sollte, ergab sich gerade für die ersten Meßzeitpunkte die Frage nach der 'subjektiven Urteilssicherheit'.
Wie wir nämlich vermuten können, haben die Lehrer zu diesen ersten Meßzeitpunkten noch wenige Informationen über die Schüler, sodaß sie sich mit ihrem Urteil nicht weit vom O-Punkt der Skala entfernen werden. Das bedeutet nun wiederum, daß der O-Punkt der Skala in diesen Fällen nicht den Mittelpunkt der Merkmalsebene darstellt, die durch die Polbezeichnung definiert ist, sondern einen Indifferenzpunkt, der gewählt wird, wenn keine Informationen vorliegen und trotzdem ein Urteil abgegeben werden soll.
In dieser Zwangssituation werden die Lehrer wahrscheinlich Ähnlichkeitsschätzungen der Eigenschaftsbegriffe vornehmen, denen dieselben Vermutungen über Eigenschaftszusammenhänge zugrundeliegen werden wie in der Stereotypen-Beurteilung. Die IPT wird also gerade zum Meßzeitpunkt 1, in einem Stadium reduzierter Informationen über die zu beurteilenden Reizpersonen gut zu rekonstruieren sein. Die errechneten Korrelationen zwischen den Eigenschaftsbegriffen werden durch die geringe Skalenausnützung nicht beeinflusst.
Steigt nun mit zunehmender Kontaktzahl die Zahl der Informationen, so wird eine steigende Urteilssicherheit (Mittelwert der Skala "subjektive Sicherheit") zu beobachten sein, damit verbunden andererseits ein Auswandern der Markierungen auf den Skalen in Richtung auf die Pole hin. Folgen wir den Annahmen der realistischen Interpretation, nach der Informationen die IPT verändern können, so ist die Wahrscheinlichkeit groß, daß die zunehmende Zahl von Informationen eine Veränderung der ursprünglich zum ersten Meßzeitpunkt fest-

gestellten Faktorenstruktur der Urteile hervorruft.

Folgen wir dagegen den Annahmen einer idealistischen Interpretation, so ist nicht anzunehmen, daß sich die zum Meßzeitpunkt 1 festgestellte Struktur der Urteile ändert. Grundsätzlich interessant ist der Wert der Skala 'subjektive Urteilssicherheit' zum ersten Meßzeitpunkt.
Die Lehrer können nach der einen abgehaltenen Schulstunde durchaus schon einige Wahrnehmungen mit den Namen einzelner Schüler verbinden. (z.B. wenn besonders auffälliges Verhalten, Aussehen usw. des Schülers zur Feststellung seines Namens geführt hat). Die Informationen können jedoch wegen der Kürze der Kontaktzeit nicht besonders differenziert sein.

PASSINI u. NORMAN (1966) vermuten, daß bereits wenige Hinweisreize aus dem äußeren Erscheinungsbild,dem Verhalten oder der Kleidung etc., bereits genügen, um die Aktivierung impliziter Persönlichkeitstheorien zu erklären.

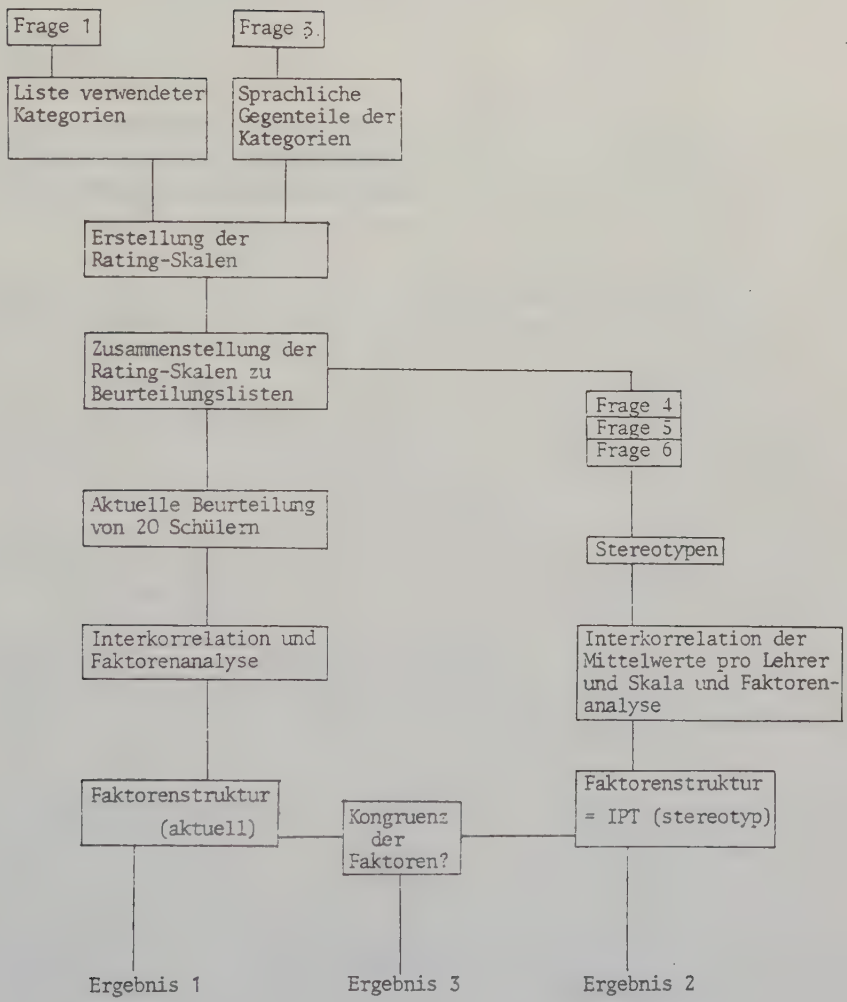
In der Beurteilungssituation, wenn der Lehrer also die Rating-Listen ausfüllt, wird bei Auftauchen dieses Namens mit einiger Wahrscheinlichkeit das in der 'life' - Situation gebildete Urteil wieder aktiviert und sich in einer vom O - Punkt der Skala verschiedenen Markierung niederschlagen. Da im Urteil aber auch Vermutungen über Eigenschaftszusammenhänge enthalten sind (s. Modell von WARR u. KNAPPER), wird nicht nur die Markierung auf einer Skala von O abweichen, sondern auf mehreren Skalen. Die Markierung auf der Skala "subjektive Urteilssicherheit" müßte dann die Sicherheit angeben, die durch die Aktivierung der vermuteten Eigenschaftszusammenhänge, also letztlich der IPT, beim Rating gewonnen wurde. Es wäre also bereits zum ersten Meßzeitpunkt mit Mittelwerten der Skala 'subjektive Sicherheit' zu rechnen, die veschieden sind von O.

4.5.4. Versuchsplan I: Vergleich der IPT aus Stereotypenbeurteilungen mit den Faktorenstrukturen aktueller Schülerbeurteilungen

Der in Abb. 4 dargestellte Ablaufplan bezieht sich auf den Vergleich der Faktorenstruktur aus den Stereotypenbeurteilungen mit der Faktorenstruktur aus aktuellen Beurteilungen noch unbekannter Schüler.

Die Stereotypen-Beurteilung wurde vor einer Lehrerkonferenz in den Ferien durchgeführt. Die aktuelle Beurteilung fand am ersten Schultag statt, an dem bereits alle Lehrer Kontakt mit den Klassen hatten, denen die später zu beurteilenden Schüler angehörten. Sofort nach der ersten Schulstunde wurden Mappen mit den Rating-Listen ausgeteilt, die Lehrer füllten die Listen jedoch zu Hause aus.
Der Begriff "noch unbekannter Schüler" wird definiert als:"20 nach Zufall ausgewählte Schüler und Schülerinnen der ersten Gymnasialklasse mit durchschnittlich 37 Schülern, nach einer Stunde Kontakt des beurteilenden Lehrers mit der Klasse."

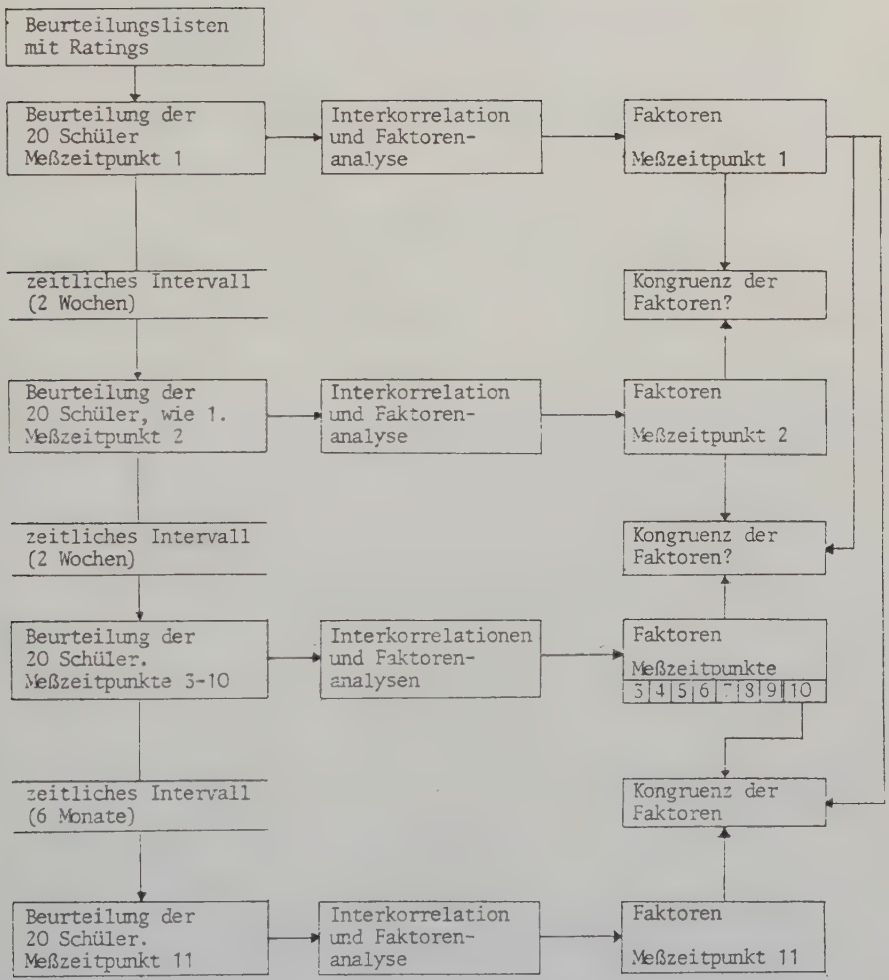
Schema des Versuchsplans I



4.5.5. Versuchsplan II - der Längsschnitt - Vergleich der Faktorenstrukturen aus den aktuellen Schülerbeurteilungen

Die 10 Beurteilungen der 20 Schüler wurden in regelmäßigen Abständen von 14 Tagen durchgeführt. An einem Tag wurden die Rating-Listen ausgeteilt, am nächsten wieder eingesammelt. Das Ausfüllen der Rating-Skalen wurde von den Lehrern zu Hause vorgenommen.
Nach dem 10. Meßzeitpunkt war das Probehalbjahr der Schüler vorbei, das alle Schüler der Beurteilten-Stichprobe bestanden. Es folgte ein Intervall von 6 Monaten, danach wurde kurz vor Schuljahrsende die Nachuntersuchung zum Zeitpunkt 11 durchgeführt.

Schema des Versuchsplans II



Ergebnis 1 : Kongruenzen der Faktoren aufeinanderfolgender Meßzeitpunkte.

Ergebnis 2 : Kongruenzen der Faktoren jedes Meßzeitpunktes mit Zeitpunkt 1.

4.5.6. Statistische Methoden

Da sich die Durchführung der Faktorenanalysen in der Querschnittsuntersuchung (Beurteilung der Stereotypen) und in der Längsschnittuntersuchung (aktuelle Schülerbeurteilungen) entsprach, sollen hier zunächst einmal Probleme der Veränderungsmessung in Längsschnittstudien dargestellt werden.

4.5.6.1. Probleme der Veränderungsmessung

Veränderung selbst kann nicht gemessen werden; sie muß erschlossen werden als Differenz zweier Zustandsmessungen.
Als einfachste Maße hierfür bieten sich an:

- Differenzmaße
  Wir gehen hier von einem Meßvorgang aus, der an der gleichen Vp wiederholt wird und so einen Eingangs- und einen Endwert liefert. Das gesuchte Maß für die Veränderung ist die Differenz zwischen beiden.

  Durch die Regression zur Mitte kommt es bei der Vorhersage von Ergebnissen der Zweitmessung aus Ergebnissen der Erstmessung zu Fehleinschätzungen besonders der Vpn, die in der Erstmessung extreme Meßwerte lieferten.

  Als Lösung dieses Problems werden in PETERMANN (1978, S.34) Differenzmaße vorgeschlagen, die entweder Differenzwerte aus standardisierten Werten der ersten und der zweiten Messung darstellen oder durch die gemeinsame Standardisierung der Werte der ersten und zweiten Messung gebildet werden.
  HELMREICH (1977, S.51 f) entwickelte zur Trendfeststellung die Möglichkeit, die Mittelwerte einer Stichprobe zu zwei Erhebungszeitpunkten bezüglich einer identischen Variablen durch t-Test der mittleren Differenzwerte zu vergleichen.
  Er bildet den 'mittleren Differenzwert' $\bar{D}$ entweder als Differenz der Summe der Eingangswerte und der Summe der Endwerte,

$$\bar{D} = \frac{SUM\ (Y) - SUM\ (X)}{N}$$

oder aus der Summe der Meßwertdifferenzen zwischen erster und zweiter Messung

$$\bar{D} = \frac{SUM\ (Y-X)}{N} \quad .$$

Bei geringerem Stichprobenumfang empfiehlt er den t-Test

$$t = \frac{\bar{D}}{\sqrt{\frac{VAR(D)}{N}}} \quad ,$$

wobei die Varianz der Differenzwerte zwischen den Messungen

$$VAR\ (D) = VAR\ (Y-X)$$
$$= VAR(Y) + VAR\ (X) - 2 * RXY * \sqrt{VAR(Y) * VAR(X)},$$

sodaß sich für t die folgende Formel ergibt:

$$t = \frac{\bar{Y} - \bar{X}}{\sqrt{\frac{VAR(Y) + VAR(X) - 2 * RXY * \sqrt{VAR(Y) * VAR(X)}}{N}}} \quad ;$$

(HELMREICH, 1977, S.52).

Diese Formel sollte allerdings nur angewandt werden, wenn die Differenzen normalverteilt sind. Falls das nicht der Fall ist, sollte der non-parametrische WILCOXON-Test gerechnet werden.

- Korrelationsmaße

Als weiteres Veränderungsmaß kann die Korrelation zwischen dem Eingangswert und der Differenz zwischen Eingangs- und Endwert errechnet werden, wobei die Differenz nicht als absoluter Betrag in die Rechnung eingehen darf, sondern mit dem richtungsweisenden Vorzeichen, ob es sich um ein Absinken oder Ansteigen des Endwertes gegenüber dem Anfangswert handelt.
Bei dieser Korrelation wird allerdings die Abhängigkeit der Messungen voneinander nicht berücksichtigt; wie wir bereits oben gesehen haben, besteht ein Zusammenhang zwischen der Höhe des Eingangswertes und der wahrscheinlichen Differenz des Eingangswertes zum Endwert:
je weiter der Eingangswert einer Vp vom Mittelwert der Stichprobe entfernt ist, desto wahrscheinlicher verändert er sich in der Zweitmessung in Richtung auf diesen Mittelwert, oder anders ausgedrückt: es besteht eine negative Korrelation zwischen Extremwerten und Differenzwerten; die Folge dieser Erscheinung ist ein negativer Beitrag dieser Werte zur Gesamtkorrelation über die ganze Stichprobe, die dadurch in der Regel unterschätzt wird. HELMREICH (1977, S.41) leitet hierzu noch Überlegungen aus dem Fehlerkonzept der klassischen Testtheorie ab, die hier jedoch nicht diskutiert werden sollen. Die Möglichkeit der Korrelation zwischen Eingangs- und Endwert wird weiter unten bei der Diskussion der Reliabilität der Ergebnisse besprochen. Insgesamt scheinen jedoch Korrelationen keine adäquaten Maße für die Feststellung von Veränderung zu sein.

- Regressionsmaße

Eine Möglichkeit, die oben erwähnten Nachteile der angesprochenen Distanzmaße auszugleichen, besteht nun darin, den Regressionseffekt zu korrigieren, indem die Regression der Zweitmessung auf die Erstmessung bestimmt wird; danach schätzt man aus den Werten der Erstmessung die Werte der Zweitmessung. Die Differenz zwischen den geschätzten und den

beobachteten Werten der Zweitmessung stellt dann die Veränderung dar. Statistisch-methodisch wird dieses Vorhaben durch eine Kovarianzanalyse realisiert, die Unterschiede zwischen zwei Gruppen hinsichtlich eines Kriteriums feststellt unter Ausschaltung der Differenz der Ausgangswerte. Das wirft zunächst die Frage der Verfügbarkeit einer Parallel-Stichprobe auf. Zum zweiten leistet bei Veränderungsmessung die Kovarianzanalyse dann keine zuverlässige Schätzung der Veränderungsbeträge, wenn zwischen den Gruppen Differenzen in den Ausgangswerten bestehen, wie PETERMANN (1978, S.57) darstellt. Derselbe Autor (a.a.O., S. 34 f) berichtet noch über ein Verfahren von CRONBACH u. FURBY (1970), das den Einfluß der Erstmessung auf die Zweitmessung über Partialkorrelation herauspartialisiert und in die Schätzung einbezieht, wobei noch die Korrelation der Veränderungsbeträge mit einer Drittvariablen diskutiert wird.
Dabei entsteht allerdings wieder das Problem, eine psychologisch bedeutsame Drittvariable theoretisch zu fundieren; je nach Wahl dieser Drittvariablen verändert sich die Zusammensetzung der Regressionsgleichung und folglich ebenfalls der durch die Regressionsgleichung definierte Veränderungsbetrag.
HELMREICH (1977, S.66 ff) diskutiert dieses Vorgehen ebenfalls und weist auf die starke Beeinträchtigung des Partialkorrelations-Koeffizienten durch Meßfehler in den Messungen der Eingangs- und Endwerte und der Drittvariablen hin, die er durch eine Minderungskorrektur der Korrelationen zu berücksichtigen versucht, die in die Partialkorrelation zwischen der Drittvariablen und den Veränderungswerten unabhängig vom Eingangswert eingehen. Das Negativum der von HELMREICH (1977, S.78) angegebenen minderungskorrigierten Korrelationsformel ist im Vergleich zur nicht-korrigierten Formel, daß nach der korrigierten Formel bereits bei einer Reliabilität der Eingangswerte von r_{tt}= .60 genau entgegengesetzte Interpretationen der gleichen Daten möglich werden, da die minderungskorrigierte Formel bei diesen Reliabilitätskoeffizienten der Eingangsdaten eine ähnlich hohe, aber

negative Korrelation zwischen Drittvariable und Veränderungswerten unabhängig von Eingangswerten produziert, wie die unkorrigierte Formel, diese jedoch mit positivem Vorzeichen. HELMREICH (1977, S.75) führt dagegen an, daß in empirischen Untersuchungen die Wahrscheinlichkeit gering sei, daß Eingangswerte einer Variablen mit einer Reliabilität in der Höhe von r_{tt} = .60 mit den Eingangswerten einer anderen Variablen eine ebenso hohe Korrelation aufweisen könnten - aber gerade das müssen wir ja nach unserer Definition der IPT als Korrelationen zwischen Eigenschaftsbegriffen (Variablen) vermuten.
Wählen wir also zur Erfassung von Veränderung den statistischen Ansatz der Partialkorrelation, wobei wir die durchaus einsehbare minderungskorrigierte Korrelation verwenden, so zerstören wir speziell auch im Fall der Überprüfung des Einflusses einer Drittvariablen auf die Veränderung zwischen Erst- und Zweitmessung unserer Variablen unser statistisches Konstrukt der IPT.
Wählen wir dagegen die unkorrigierte Korrelation, so müssen wir uns den Vorwurf gefallen lassen, Meßfehler nicht zu berücksichtigen.

Um die Methodendiskussion nicht ausufern zu lassen und uns eher einer grundlegenden Diskussion zuzuwenden, überspringen wir weitere alternative statistische Ansätze, die zum Teil auch unterschiedliche Erhebungsverfahren voraussetzen und die in PETERMANN (1978, S.40 ff) umfassend angesprochen werden; zum großen Teil lagen diese Erhebungs- und statistischen Methoden außerhalb der organisatorischen und rechentechnisch zu verwirklichenden Möglichkeiten des Autors der vorliegenden Arbeit.

Hervorzuheben wären unter diesen statistischen Methoden die faktorenanalytische T-Technik (PAWLIK, 1968, S.283), bei der die Personenstichprobe in derselben Variablen mit mehreren Messungen über die Zeit hinweg verfolgt wird; die Meßzeitpunkte werden über die Personen interkorreliert und ergeben so Situationsfaktoren für diese Variable. Negativ ist hierbei, daß die Daten trotzdem nicht ganz ausgeschöpft werden können, da bei jeder speziellen faktorenanalytischen Technik eine Bestimmungskategorie (Personen,

Situationen, Merkmale) eliminiert wird, sodaß ein Zusammenhang beim Vergleich etwa der durch T- und der R-Technik gewonnenen Faktoren nicht abgeleitet werden kann.
Eigentlich ist jedoch diese Technik nur dann angebracht, wenn man von einer Stabilität der Faktorenstruktur über die Zeit hinweg ausgehen kann, da die Meßwerte der einzelnen Zeitpunkte als unabhängige Variablen behandelt werden.
Eine Lösung dieses Problems stellt die mehrkategoriale Faktorenanalyse nach TUCKER (1963, vgl. hierzu PAWLIK, 1968, S.283) dar, bei der die Daten in einer Faktorenanalyse auf Situationsfaktoren, Merkmalsfaktoren und Personfaktoren hin analysiert werden, wobei Aussagen über den Zusammenhang dieser Faktoren gemacht werden können.
Die oben gemachten Ausführungen relativierend, seien vergleichende Leistungsüberprüfungen verschiedener Veränderungsmaße aufgeführt, die den mit z.T. erheblichem Methodenaufwand durchgeführten Datenberechnungen keine qualitativ höher zu bewertenden Leistungen zubilligen. HELMREICH (1977, S.64 ff) stellt bei der Überprüfung der Leistung dreier unterschiedlich aufwendiger Berechnungen von Veränderungsmaßen (einfachen Differenzwert, ein 'Change-Quotient', ähnlich dem 'autonomic liability score' von LACEY u. LACEY (1962) und eine weniger gebräuchliche regressionsstatistische Schätzung der "wahren" Differenzwerte nach LORD (1967) und McNEMAR (1958),) bezüglich der Veränderung auf einzelnen Skalen von Einstellungsfragebogen (HELMREICH, 1977, S. 32 f) sehr hohe Korrelationen zwischen allen drei Veränderungsmaßen fest, was für die Verwendung des ökonomischen Distanzmaßes 'Differenzwert' spricht. Über ein ähnliches Ergebnis berichtet PETERMANN (1978, S.35 f): RICHARDS (1975) und RICHARDS, KARWEIT u. PREVATT (1975) führten ein Experiment durch, in dem aufgrund computersimulierter Daten die Korrelation verschiedener Studienleistungen aus verschiedenen Determinanten (z.B. Einfluß der Schule) vorhergesagt werden sollte; die Ergebnisse zeigten eine genaue Vorhersage-

möglichkeit allein bei Verwendung des einfachen Veränderungsmaßes "Differenz".
Im nächsten Abschnitt wollen wir uns der grundlegenden testtheoretischen Problematik des Messens im Einzugsbereich der 'impliziten Persönlichkeitstheorien' zuwenden.

4.5.6.2. Validität und Reliabilität der Ergebnisse

Die Möglichkeiten einer Abschätzung der Validität der Ergebnisse stellen allerdings kein spezifisch mit Längsschnittstudien über die IPTn verbundenes Problem dar und wurden schon häufiger im Verlauf der Arbeit angesprochen.

Zunächst haben wir also das alltägliche diagnostische Vorgehen des Lehrers in der Schule, der seine Schüler permanent wahrnimmt und beurteilt. Wir versuchen, solche Urteile durch Vorlage von Eigenschaftslisten und Beurteilung der Schüler durch die Lehrer anhand dieser Eigenschaftslisten und durch Korrelieren der Eigenschaftsskalen und faktorenanalytische Bearbeitung der Korrelationen sichtbar zu machen. Während CATTEL (1957) aber die Übereinstimmung der Faktoren über verschiedene Urteiler- und Beurteilten-Gruppen noch als Beleg für die Hypothese verwendete, daß die Persönlichkeitsstruktur im Laufe der Entwicklung keinen Veränderungen unterworfen sei, wurde auf wahrnehmungspsychologischer Basis für die Erscheinung einer gewissen Faktorenkongruenz aus verschiedensten Beurteilungen ein anderes Konstrukt entworfen, die IPT, die besagt, daß die Lehrer sich bei der Reizverarbeitung weitgehend auf Vermutungen bezüglich des Zusammen-Auftretens von Eigenschaften stützen. Dieses theoretische Konstrukt IPT beinhaltet die erhebungsmethodische Implikation, daß das alltägliche Beurteilen der Lehrer anhand von Rating-Skalen mit Eigenschaftsbezeichnungen sichtbar gemacht werden kann, und die statistisch-methodische Implikation, daß Korrelationen der Skalen und Faktoren- oder Clusteranalysen Rekonstruktionen der Urteilsorganisation darstellen.

Um Belege für die Gültigkeit dieses Konstrukts als adaequate Abbildung des alltäglichen Urteilsverhaltens zu finden, wird auf verschiedene Arten versucht, Beziehungen zwischen dem Konstrukt und anderen Variablen von Urteilern und Beurteilten zu sichern.

Gerade mit diesem letzten Gesichtspunkt steuern wir die Frage der Konstruktvalidität an. Überlegungen dazu sind eher theoretischer als pragmatischer Natur. Hierzu zählt die grundsätzliche Kritik der Erhebungsmethodik und der Anwendung bestimmter statistisch-mathematischer Modelle mit ihren Implikationen wie: Annahme von Intervallskalen-Niveau der Rating-Skalen; Annahme linearer Beziehungen zwischen den verwendeten Eigenschaftsbegriffen; Annahme der Homogenität der Urteilsverteilung auf den Schätzskalen; Ansicht, die Mittelwerts- und Varianzunterschiede der Skalen seien zu eliminierende Fehler; Annahme, der Raum, in dem die Vpn die Ähnlichkeiten der Eigenschaftsbegriffe schätzen, entspreche dem euklidischen Raum mit der dort gültigen Distanz- bzw. Ähnlichkeitsmetrik, etc. (vergl.hierzu : GIGERENZER/STRUBE, 1978), wobei die Frage der Validität der Methoden zur Konstrukterfassung für die tatsächlichen Abläufe berührt wird.

Ein Ausweg besteht hier nur darin, den Gültigkeitsanspruch der Untersuchungsergebnisse auf Urteilssituationen zu beziehen, die der Untersuchungssituation sehr ähnlich sind oder aber umgekehrt, die Untersuchungssituation analog zur Urteilssituation aufzubauen - ein Unterfangen größter Schwierigkeit, wenn man bedenkt, daß die Urteile, die man durch IPTn beeinflußt vermutet, normalerweise von den Lehrern nicht einmal verbalisiert werden.

Werden sie dagegen Dritten gegenüber verbalisiert, dann stellen diese Äußerungen bereits situationsadäquate Modifikationen der Urteile dar. Wenn nun andere Untersucher in anderen Situationen mit anderen Methoden zu anderen Ergebnissen kommen, so bedeutet das ja keineswegs, daß der gemeinte

Sachverhalt 'Alltagsbeurteilung' in diesen anderen Untersuchungen eher erfaßt wird; sie stellen eigentlich nur in den Augen des Theoretikers anscheinend adäquatere Zugangsweisen zum Gemeinten dar.

Es ist jedoch zu überlegen, ob nicht im Rahmen der kriterienbezogenen Validität Hinweise für die Gültigkeit des Konstrukts für den gemeinten Sachverhalt gesammelt werden können. Die innere Validität ließe sich mit einem Vergleich von Schätzurteilen verschiedener Vpn feststellen, wobei allerdings nur die Frage der Stabilität des Konstrukts zwischen den Vpn geklärt wird, also eigentlich eine Reliabilitätsfrage.

Beim Versuch, die äußere Validität zu eruieren, müßte erst einmal ein objektiv bewertbares Kriterium gefunden werden. Benutzt man Testergebnisse als solches objektiveres Außenkriterium, so artet die Validitätsprüfung in einen Konstruktvergleich aus; die Höhe der Validität hängt dann von der theoretischen Übereinstimmung zwischen dem IPT-Konstrukt und dem Konstrukt ab, das dem Test zugrundeliegt; dieser Konstruktvergleich wird zudem gewissermaßen über die Köpfe der Urteiler hinweg vorgenommen, hängt also mit dem gemeinten Sachverhalt "Alltagsdiagnostik" gar nicht mehr zusammen. Ein solcher Konstruktvergleich scheint vorzuliegen, wenn einzelne Untersucher die Angemessenheit von Lehrerurteilen betonen (z.B. HANKE, LOHMÜLLER u. MANDL, 1978, S.44).
Natürlich könnte man die äußere Validität des Konstrukts erhöhen, indem man die Urteiler lehrt, die gleichen Indikatoren für die Zuschreibung von Persönlichkeitsmerkmalen zu verwenden, wie sie den Tests zugrundeliegen. In diesem Sinne könnte man z.B. die "Trainingskurse für die Beurteilung der Schülerpersönlichkeit" mißverstehen, die HOFER (1969, S.91) vorschlägt.

Die Erhebung der Übereinstimmungsvalidität könnte man sich als inhaltsanalytische Verarbeitung anderer expliziter Urteile (z.B. Äußerungen gegenüber den Eltern oder gegenüber der Schulbehörde)vorstellen.In diesem Sinn hat z.B. die Untersuchung von HANKE,LOHMÖLLER u.MANDL (1978) eine hohe Übereinstimmungsvalidiät erreicht, indem ihre Beurteilungsbogen als offizielles Medium der Beurteilung im Schulbereich anerkannt wurden; andererseits mußten die Autoren jedoch ihre Beurteilungskategorien von der Schulbehörde übernehmen, sodaß die Autoren selbst einräumen, daß es sich in ihrer Untersuchung weniger um die in der Alltagsdiagnostik von den Lehrern "angewandte" IPT handelt, als eher um die "offizielle" IPT der Institution Schule, die die Lehrer der Stichprobe vollendet Jahr um Jahr reproduzierten.

In der Untersuchung der Vorhersagevalidität scheint für zukünftige Untersuchungen der IPTn ein weiteres Vorankommen in der Frage der Gültigkeit des Modells für alltagsdiagnostische Prozesse zu liegen. Allerdings müßten hierzu die IPTn in den Rahmen der Untersuchung der Interaktion zwischen Urteiler und Beurteilten eingebettet werden, d.h. es müßte der gegenwärtige Wahrnehmungsprozess von Lehrern und Schülern in der Interaktion zwischen beiden möglichst unter Einbeziehung des gegenseitigen Verhaltens untersucht werden, sodaß aus den Beziehungen der beteiligten Variablen eine Vorhersage von Verhalten möglich wird.
Dieser Anspruch mag global klingen, ist aber nicht umfangreicher als der Anspruch, die Vielbedingtheit psychischer Phänomene mit den heutigen Mitteln und Methoden der Psychologie jemals umfassend in den Griff zu bekommen; schon das Messen selbst, der Zwang, zu quantifizieren verfälscht das Objekt, indem die mögliche Vielbedingtheit der Phänomene reduziert wird auf die Erfassung einiger weniger Bedingungen, weswegen die angestrebten Aussagen auch nicht genereller sein dürfen als es die Untersuchung zuläßt.
Diese Aussage hört sich trivial an, bedeutet aber für unsere Untersuchung, daß wir nur Produkte des "Rating-Skalen-Ankreuz-Verhaltens" der Lehrer untersuchten, daß ferner die Aussagen der Ergebnisse nur valide sind bzw. waren für diese spezielle, von uns vorgefundene oder konstruierte Beurteilungssituation mit diesen Stichproben von Urteilern und Beurteilten und dieser

speziellen Methodik. Es handelt sich also um eine Fallstudie zum Thema 'Implizite Persönlichkeitstheorie'.

LIENERT (1969, S.14) definiert: Reliabilität:
"Unter der Reliabilität eines Testes versteht man den Grad der Genauigkeit, mit dem er ein bestimmtes Persönlichkeits- oder Verhaltensmerkmal mißt, gleichgültig, ob er dieses Merkmal auch zu messen beansprucht...."
In unserem Fall könnten wir die Genauigkeit auf drei Ebenen der Datenerhebung bzw. -verrechnung feststellen:

1. Reliabilität der Beurteilungen
2. Reliabilität der Korrelationen
3. Reliabilität der Faktorenstruktur

Die Reliabilität von Beurteilungen wird lehrbuchmäßig durch Retest, Parallel-Test und Überprüfung der inneren Konsistenz bzw. Konsistenzanalyse geleistet.

Die Reliabilität der Korrelationen bestimmt HOFER (1969, S.72) durch den BARTLETT-Test I (s. PAWLIK, 1968, S.86 ff), der eine mehr als zufällige Abweichung der Interkorrelationsmatrix von der Null-Hypothese überprüft, nach der die vorliegende Stichprobe aus einer Population stammt, in der die verwendeten Variablen unkorreliert sind.
Die Reliabilität der Faktorenstruktur stellt HOFER (1969, S.73) fest, indem er die Stichprobe halbiert, die beiden Hälften getrennt faktorenanalysiert und die Faktoren durch orthogonale Rotation einander maximal annähert. Die Ähnlichkeitsindices werden als ein Maß der Reliabilität der Faktorenstruktur betrachtet.
Gehen wir zurück zu den Methoden der Realiabilitätsfeststellung von Beurteilungen und prüfen wir ihre Anwendbarkeit auf unsere konkrete Fragestellung der IPT unter dem Aspekt möglicher realistischer Interpretation der IPT. HOFER (1969, S.58) läßt zur Schätzung der Beurteilungs-Reliabilität 3 Lehrer je 5 Schüler aktuell im zeitlichen Abstand von einer Woche nochmals beurteilen, errechnet Retest-Reliabilitäten von r_{tt} = .65, .72, .89 und schließt daraus, daß diese Lehrer die Beurteilung ihrer Schüler "in einigermaßen zuverlässiger, z.T. sogar sehr zuverlässiger Weise vornahmen". (a.a.O., S.59).

Betrachten wir dieses Vorgehen unter unserer Fragestellung, so ist zu fragen:
Was hat sich zwischen beiden Messungen verändert, die Reliabilität des Meßinstrumentes oder das gemessene Merkmal?
Die Testtheorie geht hier von einer gewissen Konstanz des Merkmals aus, wie aus der Definition von LIENERT zu entnehmen ist. Die Bestimmung der Reliabilität per Retest scheint also für unsere Fragestellung nicht angebracht zu sein.

Die Reliabilität des Meßinstrumentes über einen Paralleltest zu überprüfen, scheint ebenso unmöglich; wir haben eine Anzahl von items, die von der Lehrergruppe als relevant für Schülerbeurteilungen angesehen wird und die nicht durch eine beliebige konstruierte andere Liste von items ersetzt werden kann.
Statt an einen Parallel-Test könnte man an einen Versuchsplan mit Kontrollgruppe denken, der bei geringeren Messungshäufigkeiten in der Kontrollgruppe eine Abschätzung des Reliabilitätsverlustes des Meßinstrumentes durch häufige Anwendung in der Versuchsgruppe gestattet.
Dabei kann allerdings auch nur die relative Abnützung des Meßinstrumentes in der Versuchsgruppe im Vergleich zur Kontrollgruppe geprüft werden, da die Kontrollgruppe ja auch mehrfach untersucht werden müßte.
Zur Erklärung von plötzlich eintretenden Veränderungen in der Versuchsgruppe durch einen Testungseffekt kann ein Versuchsplan mit Kontrollgruppe auch deswegen nur eingeschränkt herangezogen werden, weil nicht vorausgesehen werden kann, zu welchen Zeitpunkten dieses Phänomen in der VG auftritt, wenn, wie in unserem Fall, nicht schon Hypothesen aus den Ergebnissen ähnlicher Untersuchungen abgeleitet werden können.
Was bleibt, ist die Feststellung der internen Konsistenz. Hierzu müssten die Skalen jedoch homogen, also auf den gleichen Faktor zurückzuführen sein, was, wie wir wissen, nicht der Fall ist. Eine Konsistenzbestimmung durch Vergleich der Urteile von zwei Lehrern über den gleichen Schüler ist ebenfalls nicht möglich. Eine daraus errechnete geringe Reliabilität ist zu mehrdeutig: eine Diskrepanz im Verständnis

der vorgegebenen Eigenschaftsbegriffe und z.B. unterschiedliches Schülerverhalten den beiden Lehrern gegenüber sind dann als Gründe für die geringe Konsistenz nicht zu trennen.

Für die Prüfung der Reliabilität über die Differenzen zwischen Meßwerten aufeinanderfolgender Messungen hat HELMREICH (1977, S.49) eine Stabilitätsformel vorgeschlagen. Negativ daran ist, daß in dieser Formel der Parallel-Test- oder Konsistenzkoeffizient benötigt wird und daß er nur für den Vergleich von zwei Meßzeitpunkten angelegt ist.

Die Prüfung der Reliabilität der Korrelationen bzw. der Korrelationsmatrix erfolgt auch bei uns routinemäßig über den BARTLETT-Test I, der sowieso für die Sicherung des ersten extrahierten Faktors notwendig ist.

Die Prüfung der Stabilität der Faktorenstruktur stellt wieder ähnliche Probleme wie auf der Datenebene.

HELMREICH (1977, S.25,81) schlägt dafür eine Ähnlichkeitsrotation der Faktorenmatrizen aufeinander vor (FISCHER/RÖPPERT, 1964), wobei er allerdings von Meßwerten in Form von Markierungen auf jeweils einer Skala aus Einstellungsfragebogen ausgeht. Der Matrixkongruenzkoeffizient Q (GEBHARDT, 1967, S.591) stellt dann ein Maß für die Stabilität der Faktorenstruktur dar, die Faktorkongruenzkoeffizienten (WRIGLEY u. NEUHAUS, 1955; siehe: PAWLIK, 1968, S.260) geben die Stabilitäten der einzelnen Faktoren an.
Im Verlauf der wiederholten Messungen haben wir mit folgenden Quellen für eine Reliabilitätsveränderung zu rechnen:

- instrumentelle Inkonsistenz_ innerhalb des Meßinstruments
- Merkmalsinkonstanz
- Bedingungsinkonstanz
- Wiederholungseinflüsse
- Erinnerungseinflüsse

Erinnerungseinflüsse können wohl bei 2ox29 Urteilen für jeden Lehrer pro Meßzeitpunkt als Quelle für eine scheinbare Reliabilitätserhöhung vernachlässigt werden.
Soll nun aus der Untersuchung keine Aussage über die Alltagsdiagnostik der Lehrer abgeleitet werden, sondern nur über die Adaequatheit der Methoden für die Erfassung des Konstrukts IPT als "konsistente Annahme eines Eigenschaftszusammenhangs" (CLINE, 1964), (und zwar nur der vorgegebenen Eigenschaften in der Untersuchungssituation), so betrifft die "Merkmalsinkonstanz" nicht mehr die Merkmale der Schüler, sondern die Eigenschaftsbegriffe, die die Ratingskalen bezeichnen, womit die Merkmalsinkonstanz in der instrumentellen Inkonsistenz des Meßinstruments aufgeht, ebenso wie die Bedingungsinkonstanz. Die Reliabilitätsuntersuchung wird also zu einer Untersuchung der Konstruktvalidität. Hohe Kongruenzen der Faktorenstrukturen zwischen den Meßzeitpunkten zeigen dann eine Stabilität des Konstrukts an, geringe Kongruenzen eine geringe Stabilität des Konstrukts, die als Inkonsistenz der Annahmen von Eigenschaftszusammenhängen interpretiert werden könnte. Während aber bei Untersuchungen wie der von HANKE, LOHMÖLLER u. MANDL (1978) mit zeitlichen Intervallen von einem Jahr Wiederholungseinflüsse als unwahrscheinlich anzusehen sind, werden sie in unserer Untersuchung mit 2-Wochen-Intervallen schon wahrscheinlicher. Zur Erfassung dieser Wiederholungseinflüsse wäre eine nicht so häufig wie die Versuchsgruppe beurteilende Kontrollgruppe von einem gewissen Nutzen gewesen; eine eventuelle Inkongruenz der Faktorenstrukturen der Beurteilungen beider Gruppen könnte aber ebensogut von einer mangelnden Parallelität der Gruppen herrühren.

So bleibt uns nichts weiter übrig, als anzunehmen, daß der "Verbrauch"(Reliabilitätsverlust) der Skalen durch Wiederholungseinflüsse alle Skalen der Rating-Liste zu einem Meßzeitpunkt in gleicher Weise betrifft; diese Annahme ist plausibel, wenn man als Grund für den "Verbrauch" der Skalen eine sinkende Motivation der Urteiler zu weiterer Teilnahme an der Untersuchung sieht.

Wiederholungseinflüsse aus anderen Gründen, etwa in dem Sinn, daß die Urteiler durch die Darbietung sonst nicht von ihnen benutzter Beurteilungskategorien erst auf deren Existenz aufmerksam gemacht werden, müßten sich dann eigentlich in einer steigenden Differenziertheit der Urteile niederschlagen, d.h. in der Entstehung anfänglich noch nicht existierender, neuer Faktoren, die bis zum Ende der Untersuchung beibehalten werden. Diese Vermutung bedeutet für die statistische Auswertung der Urteile, daß Veränderungstendenzen einzelner Skalen verfolgt werden müssen. Durch das Errechnen von Matrixkongruenzen werden solche Tendenzen allerdings etwas verdeckt, da beim Matrizenvergleich vorrangig identische Ladungsmuster aufgesucht werden, während die absoluten Beträge der Ladungen erst bei der Errechnung der Kongruenz einzelner Faktoren verwendet werden. Wir werden daher nur Faktorkongruenzen berechnen, wie es auch die Fragestellungen der Untersuchung schon vorsehen.

Aus ebendiesem Grunde werden wir auch weder ein einheitliches Faktorenmodell schaffen, noch Faktorenanalysen höherer Ordnung rechnen; durch beide Maßnahmen werden die Faktoren uniformiert, und die Chance, Variabilität abzubilden, geringer. Während die Faktoren im nicht auf Uniformität hin behandelten "Rohzustand" als Primärfaktoren durchaus noch unterschiedlich hohe Kongruenzen zwischen den Meßzeitpunkten zeigen können, ist das bei Faktoren zweiter Ordnung unmöglich; sie werden in der Regel zwischen den Messungen "höchstkongruent" sein. Wir meinen jedoch, daß die Varianzbeträge, die ein Faktor bindet, Informationen über die Wichtigkeit dieses Faktors für die Beurteiler enthält und daß Bewegungstendenzen einzelner Skalen von einem Faktor zum anderen Rückschlüsse zulassen über die Einbettung einzelner Skalen (Eigenschaftsbegriffe) in das Gesamt der Urteilsdimensionen.

4.5.6.3. Statistische Methoden der Untersuchung

Einer Anzahl von Beeinträchtigungen,die einmal aus der Verwendung des Statistik-Programmsystems für die Sozialwissenschaften (SPSS) in der einzigen zum Zeitpunkt der Datenverarbeitung am Rechenzentrum verfügbaren Version 6 (BEUTEL et al.,1976) bestanden, zum anderen aus Restriktionen bezüglich der zugebilligten Rechenzeit, der Länge der output-files und speziell auch der permanent-files, fielen einzelne geplante, aber rechenzeitintensive Operationen zum Opfer, wie z.B. die Ähnlichkeitsrotation der Faktorenmatrizen.

Die Daten mußten also mit einfacherer statistischer Methodik analysiert werden, obwohl wir uns darüber klar sind, daß damit das Datenmaterial oft unter Niveau analysiert wurde und viele Informationen verlorengingen, da die Auswertung nicht die gesamte Struktur des Datensatzes gleichzeitig erfasst.

4.5.6.3.1. Extraktion der Faktoren

Die im Versuchsplan mit 1. und 2. bezeichneten methodischen Schritte wurden bereits im letzten Kapitel dargestellt, da sie zur Konstruktion der Meßinstrumente (Rating-Skalen) benötigt werden.
Die anhand der Rating-Skalen abgegebenen Urteile wurden zu jedem der 11 Meßzeitpunkte interkorreliert und einer Faktorenanalyse unterzogen.
Als Methode zur Extraktion gemeinsamer Faktoren wurde die Hauptachsenmethode (vgl. PAWLIK, 1968, S.60 ff) standardmäßig zu allen Meßzeitpunkten angewandt. (Hauptkomponentenmodell).
Als Kommunalitätenschätzung für die erste Iteration wurde die Diagonale mit Einsen besetzt (nach KAISER, 1960). Als Abbruch-Kriterium für die Extraktion wurde zunächst das Eigenwert-Kriterium gewählt (Kriterium S_1 der Mindestrangschätzmethode nach GUTTMAN,1954). Die extrahierten Faktoren wurden stets nur orthogonal auf Einfachstruktur der Spalten notiert (normiertes Varimax-Kriterium nach KAISER, s. PAWLIK, 1968, S.207).
Schiefwinklige Rotationen wurden zwar zusätzlich durchgeführt, werden aber wegen der Vergleichbarkeit mit der Korrespondenz-Untersuchung einerseits und einer gewissen Künstlichkeit der resultierenden Faktoren andererseits hier nur kurz angesprochen werden.
Visuelle Rotationen wurden nicht zusätzlich durchgeführt. Zur Sicherung der Faktoren wurden für die ersten extrahierten Faktoren der Bartlett-Test I gerechnet (s. PAWLIK, S.85 ff).

Bei dem Versuch einer Verallgemeinerung der Ergebnisse von Faktorenanalysen sind zwei Fehler zu berücksichtigen: einmal der Merkmalsstichprobenfehler, der aus der nur teilweisen Erfassung des Merkmalsbereiches resultiert, zum anderen der Personenstichprobenfehler, der sich aus der nur teilweisen Erfassung der Population, aus der die Stichprobe gezogen wurde, ergibt. Der Bartlett-Test I überprüft den Personenstichprobenfehler, wobei die Nullhypothese lautet: Die vorliegende Personenstichprobe ist eine Zufallsstichprobe aus der festgelegten Population, in der die verwendeten Variablen nicht korrelieren. Um H_0 zu prüfen, wird festgestellt, ob die Interkorrelationsmatrix R_X der Skalen nur zufällig oder überzufällig von der Einheitsmatrix E (Diagonalmatrix, bei der alle Diagonal-Elemente = 1, alle nicht-diagonalen Elemente = 0 sind), abweicht.
Bei df = 1/2 . m . (m-1) verteilt sich die von Bartlett vorgeschlagene Prüfgröße für kleinere Stichproben

$$-\left[n - 1 - 1/6 \cdot (2m + 5)\right] \cdot n \, |R_X|$$

nach χ^2.
(n = Anzahl der Vpn; m = Anzahl der Skalen; R_X = Determinante der Interkorrelationsmatrix $|R_X| = \lambda_1 \cdot \lambda_2 \cdot \lambda_3 \ldots \lambda_m$ = Produkt der Eigenwerte von R_X).

Ist die genannte Prüfgröße auf dem hier gewählten Signifikanzniveau von 0.05 signifikant, so weicht R_X überzufällig von E ab, d.h. die Hypothese einer Zufallsstichprobe aus der Grundgesamtheit von Personen, in der die Variablen unkorreliert sind, ist zurückzuweisen. In diesem Fall ist einmal erst die Berechtigung einer Faktorenanalyse gezeigt, zum anderen ist wenigstens der erste Faktor als statistisch gesichert zu betrachten.

Für den zweiten und alle weiteren Faktoren wurde der Bartlett-Test II gerechnet (PAWLIK, S. 87 ff).

Beim Bartlett-Test II wird aufgrund signifikanter versus zufälliger Differenzen der restlichen Eigenwerte , , ... entschieden, ob der nächste Faktor noch signifikant ist.

Ist $\chi^2 = - n \cdot \ln R^{(k)}$ $\quad \prod^{m}$

$$(R^{(k)} = (m-k)^{m-k} \cdot \frac{p = K+1}{\left(\sum_{p=K+1}^{m} \lambda_p\right)^{m-k}})$$

bei df = 1/2 . (m - k + 2) . (m - k - 1) Freiheitsgraden signifikant, so unterscheiden sich die restlichen Eigenwerte signifikant, was gemäß der oben aufgeführten Argumentation bedeutet, daß dieser nächste Faktor ebenfalls statistisch gesichert ist.

Der Bartlett-Test II wird nun für jeden weiteren Faktor wiederholt; grundsätzlich ist das solange möglich, bis der vorletzte Faktor erreicht ist. Ein weiteres einschränkendes Kriterium dieses Tests besagt jedoch, daß die oben angeführte Berechnung der Freiheitsgrade nur Gültigkeit hat, solange die bereits abgerechneten Eigenwerte noch relativ groß sind gegenüber den verbleibenden Eigenwerten. Diese Voraussetzung ist oft nicht mehr gegeben, wenn man rechnerisch beim letzten Eigenwert über 1,0 angekommen ist, weswegen dort meistens die Signifikanzberechnungen der Faktoren abgebrochen wurden, falls nicht schon vorher keine Signifikanz mehr zu erreichen war. Außerdem sagen ja auch gesicherte Faktoren inhaltlich wenig aus, wenn sie nur geringe Beträge der Merkmalsvarianz aufklären, wie es in den vorliegenden Ergebnissen oft schon bei einer relativ geringen Faktorenzahl der Fall war.

Zur groben Orientierung dient zunächst einmal eine optische Identifikation durch den Vergleich der Variablen, die über .50 auf einem Faktor laden - über die Meßzeitpunkte hinweg. Einen weiteren Anhaltspunkt zur Identifikation soll der "matching index" von CATTELL (1957) liefern.

> Der "matching index" ist ein zufallskriterisches Verfahren, das die Wahrscheinlichkeit berechnet, mit der die höchsten Ladungen bei denselben Variablen auftreten. Der kriterische Wert bei einer Signifikanz von 5% besagt, daß bei N = 29 Variablen zwei Faktoren dann als übereinstimmend zu betrachten sind, wenn 2 der 4 Variablen mit den höchsten Ladungszahlen in den beiden zu vergleichenden Matrizen gleich sind.

Für die nach den bisherigen Kriterien identischen Faktoren soll zur weiteren Identifikation der Faktorkongruenz-Koeffizient nach BURT, TUCKER und WRIGLEY & NEUHAUS (PAWLIK, 1968, S.260):

$$c_{pq} = \frac{\sum_{i=1}^{m} v_{ip} \cdot v_{iq}}{+\sqrt{\sum_{i=1}^{m} v_{ip}^2 \cdot \sum_{i=1}^{m} v_{iq}^2}}$$

berechnet werden.

> Dabei sind p und q korrespondierende Spalten der beiden Faktorenmatrizen, v je eine Faktoren-Ladungszahl. Im Gegensatz zu PEARSONS r prüft dieser Kongruenzkoeffizient die numerische Übereinstimmung der Faktorenladungen der beiden zu vergleichenden Faktoren; es kann daher nicht mit der Höhe von Korrelationskoeffizienten gerechnet werden, speziell auch deswegen nicht, weil keine Ähnlichkeitstransrotation der Matrize vorangegangen war.

4.5.6.3.4. Interpretation der extrahierten Faktoren

Der oben angesprochene BARTLETT-Test sagt zwar etwas aus über die statistische Bedeutsamkeit der Faktoren, was aber keineswegs die psychologische Bedeutsamkeit der Faktoren einschließt, der hier der Vorrang in der Diskussion eingeräumt werden soll.
Zur Interpretation der extrahierten Faktoren sollen einmal die Auflagen von PAWLIK (1968, S.265 ff) berücksichtigt werden:

1. Als Markiervariablen sollen Variablen dienen, deren Ladungen mindestens .50 betragen.
2. Als Hyperebenenvariablen werden Variablen mit Ladungen zwischen -.10 bis +.10 betrachtet.
3. Spezifische Faktoren (merkmalseigene Faktoren; "singles"), die nur eine einzige substantielle Ladung aufweisen, werden nicht zugelassen und interpretiert.
4. Spezifische Paarfaktoren ("doublets"), die nur zwei Variablen substantiell laden, werden nicht in die Interpretation einbezogen.

Um einen gewissen Ausgleich für den Ausfall der direkten Faktorenidentifikation durch Ähnlichkeitsrotation der Faktorenmatrizen zu schaffen, soll das Kriterium von FÜRNTRATT (1969) zur Bestimmung der Anzahl interpretierbarer Faktoren herangezogen werden.
FÜRNTRATT (1969) geht von folgenden Überlegungen aus:

1. Die mathematisch "richtige" Anzahl gemeinsamer Faktoren ist nicht notwendigerweise maßgeblich für die Entscheidung über die Anzahl der zu rotierenden und zu interpretierenden Faktoren. Das mathematische Kriterium (etwa das Eigenwert-Kriterium von GUTTMAN) setzt nur eine Obergrenze, aber auch darunter liegende Faktorenanzahlen können sinnvoll und interpretierbar sein.

2.) Die psycholog. interpretierbare optimale Faktorenanzahl liegt meistens unter der mathematisch "richtigen" Anzahl. Die später in der Reihenfolge extrahierten Faktoren klären in Relation zu den ersten nur noch so wenig gemeinsame Varianz, daß ein Informationsverlust durch ihre Vernachlässigung durch einen Gewinn an Determiniertheit und Interpretierbarkeit der verbleibenden Faktoren aufgefangen wird. Zu große Rotationslösungen spalten zudem oft Variablengruppen von geschlossenen Faktoren ab und rufen den Eindruck eigenständiger Dimensionen hervor, ohne aber als solche interpretierbar zu sein.

3.) Die Faktorenlösung sollte nur als Ganze oder gar nicht akzeptiert werden. Fehlerfaktoren oder unidentifizierbare Faktoren sollten, falls sie auftreten, nicht akzeptiert werden, sondern zur nächstkleineren Rotationslösung führen.

4.) Ein Faktor kann nur als interpretierbar angesehen werden, wenn er von mindestens 3 Variablen definiert wird.

5.) Eine Variable kann nur zur Interpretation eines Faktors herangezogen werden, wenn - ungeachtet ihrer Ladungshöhe - ein wirklich nennenswerter Teil ihrer Kommunalität durch den Faktor aufgeklärt wird. Dieser Fall ist gegeben, wenn das Verhältnis $a^2 : h^2$ (Quadrat der Ladungszahl : errechnete Kommunalität) $\geq .50$ ist.
Als Essenz kann man also sagen, daß eine Faktorenlösung nur dann akzeptabel ist, wenn jeder Faktor durch mindestens drei Variablen definiert ist, für die $a^2/h^2 \geq .50$, wobei der größten der sukzessiv zu überprüfenden Lösungen, die dieses Kriterium erfüllen, der Vorzug zu geben ist.

In der nun folgenden Darstellung der Ergebnisse werden die Lösungen nach dem Eigenwertkriterium und verschiedene Rotationslösungen aufgeführt.

V Ergebnisse

5.1. Stichprobenbeschreibung

5.1.1. Die Stichprobe der Beurteiler

5.1.1.1. Zusammensetzung der Stichprobe

		M	s	Variationsbreite
Alter		36,68	11,21	26 - 65
Dienstjahre		11,15	12,28	1 - 39
Geschlecht		m:16	(69,57%)	w=7 (30,43%)
Hauptfach:	Deutsch	4		
	Math.	4		
	Latein	2		
	Engl.	4		
	Physik	3		
	Biologie	2		
	Geschichte	2		
	Geographie	2		
Wochenstunden in einer Klasse:				
Lehrer in Hauptfächern:		M=5,15		
Lehrer in Nebenfächern:		M=4,92		
N		23		

Alle verteilten Merkmale wichen signifikant von der Normalverteilung ab. Diese sollte aber nicht angestrebt werden. Interessanter ist ein Vergleich mit den Verhältnissen in der Grundgesamtheit, die den GRUND- UND STRUKTURDATEN entnommen wurden:

1.) Alter

Altersgruppen der Lehrer an Gymnasien in %

	bis 30 J.	30 - 40	40 - 50	50 - 60	über 60 J.
Bundes-durchschnitt	18	44,7	21,6	11,2	4,5
Stichprobe	21,74	43,47	8,69	21,74	4,35

(auf der Basis N = 23; χ^2 3,38 bei df = 4 : n.s.)

Die Stichprobe unterscheidet sich also im Altersaufbau nicht von den Verhältnissen in der Grundgesamtheit; die Altersklasse 40-50 ist allerdings in der Stichprobe unterrepräsentiert, die Altersklasse 50-60 überrepräsentiert.
Eine Repräsentativität dieser Stichprobe ist wegen ihres Umfangs natürlich ausgeschlossen. Repräsentativität liegt nur für die Verhältnisse des Ortes vor, an dem die Studie stattfand.

2.) Geschlecht

	w	m
Bundesdurchschnitt 1974 Gymnasium in %	36,1	63,9
Durchschnitt Baden/Württ. 1974 Gymnasium in %	37,4	62,6
Stichprobe	30,43	69,57

Die Stichprobe weist weniger Lehrerinnen und mehr Lehrer auf als die Bezugsmassen.
Bemerkenswert ist, daß das Geschlechterverhältnis in dieser Stichprobe genau umgekehrt ist verglichen mit dem Verhältnis in der Korrespondenz-Arbeit von HOFER.
Unter Berücksichtigung der Rollenspezifika könnte die extrahierte Urteilsstruktur daher schon aus diesem Grund anders beschaffen sein als bei HOFER.

3.) Die Ergebnisse des FPI

Darstellung der T-Werte, getrennt nach Geschlecht

	m		w		
	M	s	M	s	Signifikanz (t-Test)
Nervosität	47,90	11,00	35,80	4,38	0.05
Aggressivität	48,27	9,43	38,62	10,78	n.s.
Depressivität	52,90	14,06	40,81	15,00	n.s.
Erregbarkeit	51,00	11,96	40,04	8,33	0.05
Geselligkeit	43,09	12,82	45,42	3,58	n.s.
Gelassenheit	45,63	9,79	46,23	8,58	n.s.
Dominanzstreben	44,90	11,25	40,05	7,2	n.s.
Gehemmtheit	53,00	9,89	49,09	5,29	n.s.
Offenheit	50,81	10,05	48,06	9,19	n.s.
Extraversion	43,72	15,16	37,03	6,24	n.s.
Emot.Labilität	52,00	12,34	40,87	12,87	n.s.
Maskulinität	45,10	8,97	56,01	8,94	0.05

Andere Differenzierungen der Ergebnisse sollen vorerst nicht durchgeführt werden.

5.1.2. Die Stichprobe der Beurteilten

Klasse	Knaben	Mädchen	Ges.	M Alter
1	12	8	20	10; 9
2	14	6	20	10;11
3	14	6	20	10;11
4	3	17	20	10; 9
5	8	12	20	10;11
6	12	8	20	10;11
7	14	6	20	10;11
	= 77 (55%)	= 63 (45%)	=140	M=10;10

Ein Vergleich mit den GRUND-UND STRUKTURDATEN ergab, daß di
Verteilung in der Gesamtstichprobe nicht von der Verteilung
in der Grundgesamtheit abweicht.

	Knaben	Mädchen
Bundesdurchschnitt Gymnasien	53,4 %	46,6 %
Stichprobe	55 %	45 %

(χ^2= 0,1 bei df = 1 n.s.)

5.2. Faktorenanalyse der Stereotypenbeurteilungen

5.2.1. Die Faktorenstruktur der Stereotypenbeurteilungen

Die Beurteilungen des "guten Schülers", des "schlechten Schülers" und "des" Schülers der 1. Gymnasialklasse waren für jede Skala und jeden Lehrer gemittelt worden. Das Resultat stellt die 'Stereotypen' (CRONBACH,1955) bzw. die 'subjektiven Erwartungswerte' (COHEN, 1969) der Beurteiler dar. Die Mittelwerte wurden über alle Lehrer hinweg korreliert und einer Faktorenanalyse unterzogen, wobei die Extraktion nach dem Eigenwertkriterium erfolgte.

Die Eigenwerte legten eine 5-Faktoren-Lösung nahe, wobei allerdings der fünfte Eigenwert mit $\lambda = 1{,}05$ sehr knapp über dem Mindestkriterium lag; der Eigenwertverlauf zeigte bei $\lambda = 1$ keine eindeutige Diskontinuität, weswegen wir mit Vorbehalt fünf Faktoren extrahieren.

Wegen der Unübersichtlichkeit der Originalmatrix (im Anhang) werden in Tabelle 5 nur alle Ladungen über .50 sowie die Prüfgröße $p = a^2/h^2$ als Interpretationshilfe angegeben.

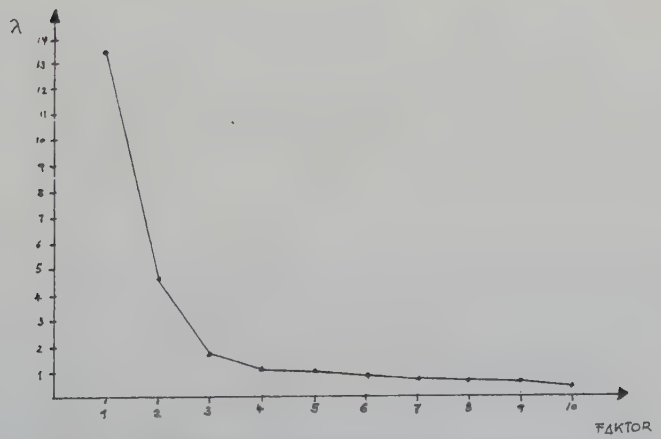


Abb.4 : Eigenwertdiagramm "Stereotypen"

Tab.4: Faktorenstruktur der Stereotypen
- Eigenwertlösung -

		FAKTOR									
Kat.	Skala (pos.Pol)	1		2		3		4		5	
Extr. Varianz (%)		64,3		19,6		7,8		4,6		3,8	
Variable		a	P	a	P	a	P	a	P	a	P
1	aufgeweckt	74	76								
2	ehrlich			61	60						
3	ehrgeizig	57	54								
4	sprachl. begabt										
5	eig.Mitdenken	78	80								
6	fleißig	56	85	76	85						
7	gt.Lehrerkontakt					68	70				
8	gt. Mitarbeit	65	74								
9	hübsch					65	45				
10	höflich			76	73						
11	intelligent	81	85								
12	kontaktfreudig					75	86				
13	konzentriert			79	89						
14	kreativ	68	66								
15	ausdauernd	53	72	66	73						
16	ordentlich			74	77						
17	paßt s.Klasse an			54	54						
18	pos.Einst.Schule			69	76						
19	akzeleriert							58	57		
20	sensibel									67	92
21	gt.häusl.Milieu										
22	unterrichtsstör.			75	67						
23	unterfordert	64	76					54	76		
24	begabt	81	82								
25	selbständig	51	63								
26	sympathisch					60	62				
27	off.Charakter										
28	vital							62	65		
29	zuverlässig			80	87						

Faktorenstruktur der Stereotypen

5.2.2. Interpretation der Faktoren

FAKTOR I (64,3% der extrahierten Varianz).
Auf ihm laden die Skalen "intelligent-", "begabt-", "eigenes Mitdenken-", "aufgeweckt-", "kreativ-", "gute Mitarbeit-", "unterfordert-"; die Skalen "selbständig-" und "ehrgeizig-" sind nur schwach repräsentiert.
Grundsätzlich weisen die geladenen Variablen auf einen Intelligenzfaktor hin, jedoch mit einer Aktivitätskomponenten (aufgeweckt, gute Mitarbeit) und einer Komponenten intrinsischer Motivation (ehrgeizig).
Da er aber primär geistige Potenz anzeigt, soll er "Geistige Regsamkeit" genannt werden.

FAKTOR II (19,6% der extrahierten Varianz)
Er lädt die Skalen: "konzentriert-", "zuverlässig-", "fleißig-", "ordentlich-", "positive Einstellung zur Schule-", "höflich-", "ausdauernd-", "stört den Unterricht nicht-", "ehrlich-", "paßt sich der Klasse an-".
Dieser Faktor faßt Eigenschaften der Arbeitsdisziplin, übertritt aber mit 'höflich-' und 'ehrlich-' diesen Bereich; wie in Faktor I eine Motivationskomponente enthalten war, ist auch hier mit 'positive Einstellung zu Schule und Lernen-'ein "im" Schüler selbst liegender begründender Aspekt angesprochen. Faktor II wird Schuldisziplin genannt.

FAKTOR III (7,8% der extrahierten Varianz)
Die Skalen "kontaktfreudig-", "guter Lehrerkontakt-", "sympathisch-", "hübsch-" werden von diesem Faktor wesentlich geladen, was diesen als Kontakt-Faktor kennzeichnet.
Interessant sind die Ladungen auf 'sympathisch-' und 'hübsch-', den Skalen, die für die Sammlung der affektiven Urteilskomponenten eingeführt worden waren.

Die Kontaktfreudigkeit eines Schülers bezüglich seiner Mitschüler und des Lehrers scheint mit emotional-affektiven Komponenten besetzt zu sein. Es kann jedoch kein gerichteter Kausalzusammenhang zwischen dem Sozialkontakt der Schüler und affektiven Zuschreibungen angegeben werden. Die beiden letzteren Skalen weisen nur mittelhohe Kommunalitäten auf, sie kommen als 'zentrale Eigenschaften' nicht in Frage.

FAKTOR IV (4,6% der extrahierten Varianz)
Bei diesem Faktor scheint es sich um einen spezifischen Paarfaktor, im Sinne PAWLIKs (1968, S.265) zu handeln. Der Faktor erklärt keinen überdurchschnittlich nennenswerten Teil der Skalen -Kommunalitäten auf; die Skalen "vital-" und "akzeleriert-" heben sich vorwiegend von einer Hyberebene aus "Disziplin"-Variablen ab, weisen also auf eine gewisse auffällige Lebendigkeit hin, auf Vitalität, wie dieser Faktor unter Vorbehalt benannt werden soll.

FAKTOR V (3,8% der extrahierten Varianz)
Dieser spezifische Faktor mit nur einer einzigen substantiellen Ladung auf der Skala "gutes vs schlechtes häusliches Milieu", ein Faktor des Sozialstatus, der nur durch die Skala mit der höchsten Kommunalität (zentrale Eigenschaft) gebildet wird, ist mit Sicherheit ein Methodenfaktor, bedingt durch die Erhebungsmethode. Erinnern wir uns, daß wir in der Vorgabe der zu beurteilenden Stereotypen mit zwei dieser Stereotypen, "der gute" und "der schlechte Schüler", das besondere Augenmerk der Lehrer auf die Leistungsdimension lenkten. HÖHN (1967, S.74 ff) stellte bei den Lehrern ihrer Untersuchung einen Vorrang der 'Milieuhypothese' zur Erklärung der Schulleistung fest; SCHWARZER (1976, S.79) weist ebenfalls auf die Relevanz der Sozialschicht der Beurteilten für das Lehrerurteil hin.

Aller Wahrscheinlichkeit nach haben wir diesen Faktor also durch die Vorgabe überwiegend leistungsorientierter Stereotypen zur Beurteilung erzeugt.
Folglich können wir die Reproduktion dieses methodenspezifischen Faktors aus Urteilen ohne diese Vorgaben (z.B. aktuelle Urteile) nicht mehr erwarten.

Durch die Übereinstimmung der durch die einzelnen Faktoren geladenen Skalen mit ebendiesen Skalen in der Untersuchung von HOFER (1969, S.75 f) können wir eine Ähnlichkeit unseres Faktors "Schuldisziplin" mit seinem Faktor "Zuverlässigkeit vs Unzuverlässigkeit", unsere "Geistige Regsamkeit" mit seinem Faktor "Begabung" feststellen.

Ein solcher zahlenmäßiger und inhaltlicher Vergleich zwischen zwei Untersuchungen ist jedoch recht riskant, da Art und Zahl der verwendeten Variablen die schließlich extrahierten Faktoren weitestgehend determinieren.

Um den Leser nicht zu irritieren, sei hier gesagt, daß wir der sprachlichen Benennung der Faktoren keinerlei Bedeutung zumessen. Sie soll nur der schnelleren Identifikation des Faktors bei späteren Vergleichen dienen und ungefähr die Tendenz des Inhalts an Skalen andeuten, die von diesem Faktor geladen werden. Die Fragestellung unserer Untersuchung bringt es aber mit sich, daß wir an den unter einem Faktor zusammengefaßten Skalenbezeichnungen (Eigenschaftsbezeichnungen) durchaus ein Interesse haben, denn die Faktorisierung läßt ja einen Zusammenhang dieser Eigenschaftsbezeichnungen, ein "Ähnlichkeitserleben" der Urteiler vermuten.

Wenden wir uns jetzt also lieber unserer zweiten Fragestellung zu, indem wir die Faktoren aus aktuellen Beurteilungen noch unbekannter Schüler extrahieren und sie mit den Faktoren aus den Stereotypenbeurteilungen vergleichen.

5.3. Faktorenanalyse aktueller Urteile über "unbekannte Schüler"

5.3.1. Faktorenstruktur der Urteile

Die Tabelle auf der nächsten Seite zeigt die nach Eigenwert-Kriterium extrahierte und varimax-rotierte Faktorenmatrix. (Korrelationstabelle und unrotierte Matrix im Anhang).

Da die Originalmatrix doch etwas unübersichtlich ist, wurden alle Ladungen über .50 (Markiervariablen) gesondert und die Hyperebenenvariablen (+.10 - -.10) in der Tabelle auf der übernächsten Seite aufgeführt. Zusätzlich wurde die als Interpretationshilfe errechnete Prüfgröße $P = a^2/h^2$ von FÜRNTRATT (siehe Abschnitt "Statistische Methoden") ebenfalls dieser Tabelle aufgeführt, sofern sie die Größe von .50 überschritt.

Die Eigenwerte zeigen eine 5-Faktorenlösung an; allerdings ist an der Stelle $\lambda = 1$ nicht die von GUTTMAN (s. PAWLIK, S.1 als zweites Kriterium für die Rangschätzung aufgeführte Diskontinuität festzustellen, weswegen wir den "wahren" Rang der Matrix mit Vorbehalt als 5 angeben.

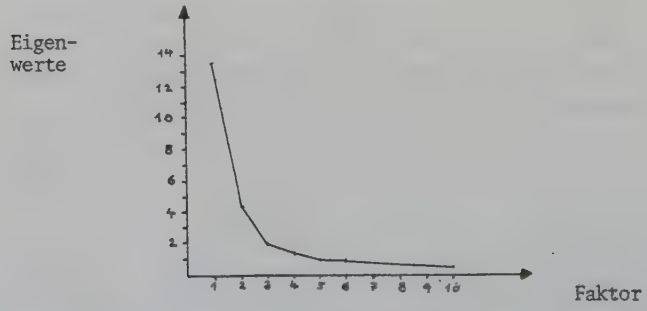


Abb 5 : Eigenwertdiagramm für die Faktorenanalyse I

Tab.5: Faktormatrix "relativ unbekannte Schüler"
- Eigenwertlösung -

	FAKTOR											
	1		2		3		4		5			
Kat. Skala (pos.Pol)	64,2		19,7		7,7		4,7		3,7			
Extr. Varianz (%)	a	P	a	P	a	P	a	P	a	P		
Variable												
1 aufgeweckt	76	74	09	11					05	11		
2 ehrlich			61	57								
3 ehrgeizig	65	71			04	11	03	11	04	11		
4 sprachl. begabt			09	11					65	67		
5 eig.Mitdenken	80	10										
6 fleißig	61	-	66	51	09	H			10	H		
7 gt.Lehrerkontakt					76	77	00	11				
8 gt. Mitarbeit	68	61					07	11				
9 hübsch	00	11	01	11	71	85						
10 höflich	06	11	84	86			07	11				
11 intelligent	84	83							07	11		
12 kontaktfreudig			03	11	75	65			06	11		
13 konzentriert			73	63	10	11						
14 kreativ	73	75							01	11		
15 ausdauernd	60	-	61	50	08	11	10	11	01	11		
16 ordentlich			75	70			09	11				
17 paßt s.Klasse an	01	11	65	61	51	-						
18 pos.Einst.Schule			70	63			04	11				
19 akzeleriert			08	11			68	67				
20 sensibel							-68	63				
21 gt.häusl.Milieu			55	-								
22 unterrichtsstör.	02	11	75	73					05	11		
23 unterfordert	66	54			00	11	55	-				
24 begabt	82	81										
25 selbständig	54	-					53	-	07	11		
26 sympathisch							04	11	55	-		
27 off.Charakter					66	65			04	11		
28 vital			07	11			72	73				
29 zuverlässig			76	67	07	11						

Faktorenladungen nach Eigenwert-Kriterium
a = Ladung; P = a^2/h^2; H - Variable der Hyperebene

Die Faktoren sind jedoch eindeutig definiert, keine Variable wird bedeutsam von mehreren Faktoren aufgeladen.

5.3.2. Interpretation der Faktoren

(Um den Leser nicht mit einer endlosen Aufzählung von Gegensatzpaaren zu ermüden, wird im folgenden oft nur der positive Pol der Skala aufgeführt, mit einem Bindestrich, z.B. "sensibel-" - gemeint ist aber die ganze Skala

FAKTOR I : (64,2 % der extrahierten Varianz)
Auf ihm laden die Eigenschaftsskalen "eigenes Mitdenken - kein eigenes Mitdenken", "intelligent - nicht intelligent", "begabt - unbegabt", "kreativ - nicht kreativ", "aufgeweckt - träge", "ehrgeizig - nicht ehrgeizig", "gute Mitarbeit - schlechte Mitarbeit", "unterfordert - überfordert".

In 5 von 8 Skalen entspricht dieser Faktor dem, den HOFER als "geistige Regsamkeit vs. geistige Trägheit" bezeichnet hat, und der hier wegen der Übereinstimmungen "geistige Regsamkeit" benannt werden soll. Auch die Skala "unterfordert - überfordert" kann unter diese Benennung subsumiert werden, einzig die Skala "ehrgeizig -" weist auf eine Komponente der Selbstdisziplin oder Eigenmotiviertheit hin, die bei HOFER's Volksschullehrern nicht hier auftrat.

FAKTOR II: (19,7 % der extrahierten Varianz)
Folgende Skalen laden auf Faktor 2:
"fleißig - faul", "zuverlässig - unzuverlässig", "ordentlich - schlampig", "konzentriert - unkonzentriert", "ausdauernd - nicht ausdauernd", "stört nie - stört häufig", "positive Einstellung zu Schule und Lernen - negative Einstellung", "höflich - frech", "ehrlich - unehrlich".

Auch hier ergeben sich wieder Übereinstimmungen mit den Ergebnissen von HOFER. Dieser Faktor soll daher "Schuldisziplin" genannt werden. Wie bei Faktor I ist eine Variable der Eigenmotiviertheit "positive Einstellung zu Schule und Lernen -" enthalten. Mit "höflich - frech" und "ehrlich - unehrlich" geht dieser Faktor über die reine Arbeitsdisziplin hinaus und beinhaltet ebenfalls diszipliniertes Verhalten gegenüber dem Lehrer als Person und Autorität.

Faktor III: (7,7 % der extrahierten Varianz)
Er wird von folgenden Markiervariablen definiert:
"kontaktfreudig - kontaktarm", "guter Lehrerkontakt - schlechter Lehrerkontakt", ("paßt sich Klasse an - paßt sich nicht an"), "offener Charakter - verschlossener Charakter", "hübsch - nicht hübsch".
Hyperebene: "ehrgeizig -", "konzentriert -", "ausdauernd -", "zuverlässig -", "unterfordert -", "fleißig - ".

Dieser Faktor, der mit "Kontaktoffenheit" umschrieben werden soll, ist ebenfalls in der Korrespondenz-Untersuchung extrahiert worden.
In der Hyperebene tauchen größere Teile des "Disziplin"-Faktors II auf, gegen den sich dieser Faktor abhebt.
Hier erscheint erstmalig eine der künstlich eingeführten Skalen: "hübsch -", die wiederum in der Hyperebene des "Disziplin"-Faktors zu finden ist.
Theoretisch könnte diese Konstellation bedeuten, daß die Lehrer einem kontaktfreudigen Schüler, speziell auch wenn er mit ihnen Kontakt aufnimmt, eher ein nettes Äußeres zuschreiben, ihn eher nett aussehend finden.

Faktor IV : (4,7 % der extrahierten Varianz)
Darauf laden: "akzeleriert - retardiert", "vital - schwächlich", die sich von dem Hintergrund "gute Mitarbeit -", "höflich -", "ausdauernd -", "ordentlich -", "positive Einstellung zu Schule und Lernen -", "sympathisch -" (Null-Ladungen) abheben.

Eine beachtenswerte negative Ladung weist die Skala "sensibel - seelisch robust" auf; recht hoch, aber nicht bedeutsam auch "stört nie - stört häufig".

Dieser Faktor stellt eigentlich ein 'doublet' dar, ist aber recht gut interpretierbar.

Er ist gekennzeichnet durch die Formierung des eindeutigen Gegensatzes "vital-", sensibel-".
Man könnte diesen Faktor "Empfindsamkeit vs. Robustheit" nennen, wie bei HOFER, er geht jedoch durch das oben Gesagte darüber hinaus - etwas in Richtung auf "Auffälligkeit - Unauffälligkeit" und müßte außerdem umgepolt werden in "Robustheit". Er soll "psycho-physische Konstitution" genannt werden.

FAKTOR V : (3,7 % der extrahierten Varianz)
Er lädt nur eine einzige Markiervariable für die Interpretation bedeutsam: "sprachlich begabt - mathematisch begabt". Dieser Faktor "spezielle Begabung" ist mit Sicherheit ein 'instrument factor', der letzlich aus der Stichprobenzusammensetzung resultiert. Die Stichprobe setzte sich zum großen Teil aus Lehrern der Fächer Deutsch, Fremdsprachen und Mathematik zusammen (ca. 61 %), bei HOFER nur 42 %, die ihre speziellen Ansichten über beurteilenswerte Schülereigenschaften in den Nennungen auf die Fragen 1 und 2 einbrachten. Da die Rating-Listen aber aufgrund der Häufigkeiten der Nennungen zusammengestellt wurden, ist diese fachspezifische Skala in der Rating-Liste "überrepräsentiert" und scheint einen eigenen "Sprachlehrer - Mathematiklehrer" - Faktor zu konstruieren, der aber so wenig Varianz aufklärt, daß er vernachlässigt werden kann.
Wäre der Varianzbetrag dieses Faktors allerdings wesentlich höher, könnte der Faktor evtl. ein Indiz für eine fachspezifisch Sichtweise von Untergruppen der Stichprobe sein.
Erstaunlich ist, daß er die Skala "begabt - unbegabt" nicht auflädt und daß "intelligent -" zu seiner Hyperebene gehört.

Die ersten extrahierten Faktoren zeigen große Ähnlichkeiten mit denen von HOFER (definiert nach den von den Faktoren geladenen Skalen, nicht nach den von HOFER angeglichenen Benennungen).

Die Beurteilungsdimensionen scheinen jedoch bei unserer Stichprobe von Gymnasiallehrern eine andere Wichtigkeits-Reihenfolge zu haben als bei HOFERs Stichprobe von Volksschullehrern.

Faktor-rang	Volksschullehrer HOFER (1969,S.60f)	Gymnasiallehrer
1.	Disziplinertheit -	Geistige Regsamkeit
2.	Geistige Regsamkeit -	Schuldisziplin
3.	Bescheidene Zurückhaltung -	Kontaktoffenheit
4.	soziale Aufgeschlossenheit -	Psycho.-phys.Konst.
5.	Empfindsamkeit -	(spezielle Begabung)

Während bei den Volksschullehrern die Disziplin die Leitdimension der Beurteilung darstellt, ist es bei den Gymnasiallehrern die 'geistige Regsamkeit'.
Diese Erscheinung könnte einmal in der Schulform begründet sein; leider geht aus HOFERs Untersuchung nicht das Alter der aktuell beurteilten Schüler oder die Klassenstufe hervor; aus den angegebenen Hauptfächern der Lehrer (falls sie in diesen eingesetzt wurden und nicht fachfremd beschäftigt wurden) läßt sich aber erschließen, daß wenigstens ein Viertel der Vpn in höheren Klassen lehrt als der 4. .Daß in diesen Altersstufen und in der Volksschule allgemein (Besuch ist Pflicht!) andere, vorwiegend disziplinarische Probleme im Vordergrund stehen als im Probehalbjahr des Gymnasiums, dürfte einsehbar sein. Ebenso, daß das Rollenverständnis des Lehrers einer Schulart von dieser im Vordergrund stehenden Problematik geprägt wird und somit auch andere Beurteilungsstrukturen resultieren. Die extrahierten Faktoren entsprechen ihrem Gehalt nach weitgehend denen von HOFER, trotzdem kann die Faktorenlösung nach dem Eigenwertkriterium nicht befriedigen, da einmal schon der 3. extrahierte Faktor nur noch sehr wenig zur Aufklärung der Varianz beiträgt, zudem ist nach dem Bartlett-Test nur der erste Faktor höhersignifikant, der zweite mit einem $z = 3{,}26$ bei $df = 405$ signifikant, der dritte mit einem $z = 1{,}72$ bei $df = 378$ auf dem 5%-Niveau signifikant, der vierte ergibt bereits einen negativen z-Wert, ist also nicht signifikant.

Der Verdacht liegt also nahe, daß nach dem Eigenwertkriterium zu viele Faktoren extrahiert wurden. Sehen wir uns daher die 4- und die 3-Faktorenlösung der gleichen Korrelationsmatrix an. (Eine "bereinigte"Darstellung findet sich auf den nächsten beiden Seiten, die Originalmatrizen im Anhang).

Bemerkenswert ist hier, daß bei beiden Lösungen einmal die Faktoren der Eigenwert-Lösung recht genau wieder erscheinen, zum anderen, daß jetzt eine Veränderung der Stellenwerte der Dimensionen eingetreten ist. Während in der oben dargestellten Lösung der Faktor "Geistige Regsamkeit" Leitdimension war, ist es jetzt der Faktor "Schuldisziplin", während die "Geistige Regsamkeit" auf den zweiten Wichtigkeits-Rangplatz gerückt ist.

Die kleinere Rotationslösung beinhaltet stets die größere. Die extrahierte Varianz muß jetzt aber durch 4 (3) Faktoren erklärt werden, was dazu führt, daß die Skalen, die in der größeren Lösung einen eigenen Faktor gebildet haben, auf die jetzt geringen Faktorenzahl verteilt werden, durch diese geladen werden. Dadurch kann es allerdings zu einer Komplexität der Faktoren kommen, die nicht mehr interpretiert werden kann, wenn man die Zahl der zu extrahierenden Faktoren zu gering ansetzt. Es ist also die optimale Faktorenlösung zu suchen, die optimal zwischen den folgenden konträren Kriterien liegt:

1. je mehr Variablen einem Faktor zugeordnet werden können (Zahl der $a^2/h^2 \geq 50$), desto besser ist diese Faktorenlösung. Je kleiner die Lösung, desto eher wird dieses Kriterium erfüllt.

2. Es ist die größte Lösung anzustreben, bei der alle Faktoren durch wenigstens 3 Variablen definiert sind, und zudem einen nennenswerten Anteil der Kommunalität dieser Variablen aufklären (s. FÖRNTRATT, 1969, S.67 f.).

3. Jeder Faktor muß eindeutig interpretierbar sein.

Tab.6: Faktormatrix "relativ unbekannte Schüler"
4-Faktoren-Lösung

Kat.	Skala (pos.Pol)	FAKTOR 1		2		3		4	
Extr. Varianz (%)		67,5		20,5		8,0		4,6	
Variable		a	P	a	P	a	P	a	P
1	aufgeweckt			71	68				
2	ehrlich	61	52						
3	ehrgeizig			55	58				
4	sprachl. begabt					46	68		
5	eig.Mitdenken			77	75				
6	fleißig	70	61	56	36				
7	gt.Lehrerkontakt					69	78		
8	gt. Mitarbeit			64	56				
9	hübsch					65	96		
10	höflich	81	94						
11	intelligent			81	77				
12	kontaktfreudig					76	67		
13	konzentriert	74	70						
14	kreativ			66	68				
15	ausdauernd	66	55	53	-				
16	ordentlich	75	69						
17	paßt s.Klasse an	53	55						
18	pos.Einst.Schule	70	60						
19	akzeleriert							65	75
20	sensibel							50	46
21	gt.häusl.Milieu	54	-						
22	unterrichtsstör.	73	90						
23	unterfordert			66	61				
24	begabt			81	79				
25	selbständig			52	-				
26	sympathisch					55	63		
27	off.Charakter					62	64		
28	vital							62	61
29	zuverlässig	79	77						

Faktorenladungen der 4-Faktorenlösung

Tab.7: Faktorenmatrix "relativ unbekannte Schüler"
3-Faktoren-Lösung

Kat.	Skala (pos.Pol)	FAKTOR 1		2		3	
Extr. Varianz (%)		70,4		21,4		8,2	
Variable		a	P	a	P	a	P
1	aufgeweckt			71	73		
2	ehrlich	69	81				
3	ehrgeizig	54	54				
4	sprachl. begabt					42	65
5	eig.Mitdenken			73	64		
6	fleißig	82	81				
7	gt.Lehrerkontakt					65	74
8	gt. Mitarbeit	50	-	57	-		
9	hübsch					65	98
10	höflich	69	85				
11	intelligent			82	79		
12	kontaktfreudig					84	81
13	konzentriert	74	72				
14	kreativ			69	74		
15	ausdauernd	73	73				
16	ordentlich	75	74				
17	paßt s.Klasse an					48	50
18	pos.Einst.Schule	73	69				
19	akzeleriert			59	85		
20	sensibel	59	68				
21	gt.häusl.Milieu	47	42				
22	unterrichtsstör.	73	86				
23	unterfordert			80	93		
24	begabt			78	77		
25	selbständig			67	79		
26	sympathisch					52	59
27	off.Charakter					65	69
28	vital					54	50
29	zuverlässig	87	87				

Faktorenladungen der 3-Faktorenlösung

Ein Vergleich der drei Faktorenlösungen nach diesen Kriterien soll eine Entscheidungshilfe sein.

Lösung	Anzahl $a^2/h^2 \geq 50$ je Faktor	aufgeklärte Kommunalität
5-Faktorenlösung	8 10 4 3 1 = 26	83 %
4-Faktorenlösung	10 9 5 3 = 27	97,5%
3-Faktorenlösung	11 8 6 = 25	97,7%

Die 5-Faktorenlösung, die nach dem Eigenwertkriterium errechnet wird, klärt zu wenig Kommunalität auf und enthält zudem einen 'instrument factor' ("spezielle Begabung"), sollte also als ganze zurückgewiesen werden.
Optimal ist die 4-Faktorenlösung; sie entspricht allen drei Kriterien (bis auf Faktor 3, Skala 20 "sensibel-").
Auch die 3-Faktorenlösung entspricht formal den aufgeführten Kriterien, ist aber weniger gut interpretierbar: unter dem Faktor "Schuldisziplin" sind jetzt die Skalen "gutes häusl. Milieu-" und "sensibel-" zu interpretieren, unter dem Faktor "Geistige Regsamkeit" zusätzlich die Skala "akzeleriert-", die sich doch auf ein physisches Merkmal bezieht, der Faktor "Kontaktoffenheit" lädt jetzt zusätzlich die Skala "sprachlich begabt-". Diese Lösung soll zwar nicht ganz abgelehnt werden, aber wenigstens für diesen Teil der Untersuchung, in dem es auf den Vergleich der Faktoreninhalte mit denen der Korrespondenzarbeit ankommt, soll nicht auf einer zwar eigenständigen, aber in den Kontext der korrespondierenden Literatur nicht einzuordnenden Faktorenlösung beharrt werden, die nicht einmal besonders gut zu interpretieren ist.

In dieser Arbeit soll also vorerst die 4-Faktorenlösung vertreten werden:

FAKTOR I :(bindet 67 % der extrahierten Varianz)
Er entspricht dem oben beschriebenen Faktor "Schuldisziplin" sowohl in den auf ihm ladenden Skalen als auch in den Ladungshöhen. Hinzugekommen ist Skala "paßt sich der Klasse an-", sodaß der Faktor jetzt neben Arbeitsdisziplin und disziplviniertem Verhalten gegen den Lehrer auch noch Gruppendisziplin erfaßt.

FAKTOR II :(20,5 % der extrahierten Varianz)
Auch dieser Faktor gleicht dem Faktor "Geistige Regsamkeit" der 5-Faktorenlösung.

FAKTOR III:(8 % der extrahierten Varianz)
Hier ergeben sich die ersten Veränderungen zur 5er-Lösung. Neu hinzugekommen sind die Skalen "sympathisch-" und leider auch "sprachlich begabt-". Damit hat sich der emotionale Aspekt komplettiert, der mit der Ladung dieses Kontaktfaktors auf "hübsch-" in der 5er-Lösung begonnen hatte.

Eine Empfindung emotionaler Zuwendung, so könnte man interpretieren, ist also für den Lehrer mit dem Kontakt zu einem Schüler verbunden. Allerdings können "hübsch-" und "sympathisch-" nicht als Anzeichen für die von ASCH (1946), KELLY (1950) und HAYS (1958) festgestellten Dimensionen "warm-kalt" in Personen-Beurteilungen gewertet werden, da diese Skalen in keiner der Lösungen die höchste Kommunalität aufweisen, also als "zentrale Variablen"nicht in Frage kommen. Über die Ladung der Skala "sprachlich begabt-" könnten weitere Vermutungen angestellt werden, von der rein methodenspezifischen Erklärung (s.o.: Spezielle Begabung), die noch am wahrscheinlichsten und am wenigsten spekulativ ist, über die Sprachbedingtheit des Schüler - Lehrer - Kontaktes bis zum Stereotyp "extravertiert".

Dieser Faktor ist wegen der stärkeren emotionalen Beteiligung "Zuwendung" genannt worden, womit die Kontakt- und Emotionalsbeteiligung zum Ausdruck gebracht werden sollen.

FAKTOR IV : (4,6 % der extrahierten Varianz)
Dieser Faktor blieb als "psycho-physische Konstitution" erhalten.

5.4. Kongruenz der Faktoren aus Stereotypenbeurteilungen und aktuellen Beurteilungen relativ unbekannter Schüler

Der Vergleich der beiden Faktorenmatrizen erbrachte folgendes Ergebnis:

Tab.8 : Kongruenzkoeffizienten C der Faktoren aus den 'subjektiven Erwartungswerten' und aus aktuellen Beurteilungen'relativ unbekannter' Schüler

FAKTOR	C
I	.99
II	.99
III	.96
IV	.93
V	.66

Als zwischen den verschiedenen Erhebungsmethoden stabil zeigten sich die ersten vier Faktoren. Im Gegensatz zur Vorbildarbeit konnte der jeweils fünfte Faktor als Methodenfaktor identifiziert werden, der nicht mehr reproduziert wurde.
Die Kongruenzkoeffizienten lagen in unserer Untersuchung jedoch wesentlich höher als bei HOFER (1969, S.82); bei ihm betrugen sie nur 0.92, 0.87, 0.68, 0.78, 0.73, was ihm jedoch schon genügt, um zusammenzufassen, "daß die an zwei verschiedenen Methoden und an zwei verschiedenen Lehrergruppen erhaltenen impliziten Persönlichkeitstheorien der Lehrer die Beurteilungen ihrer eigenen Schüler in hohem Ausmaß determinieren". (a.a.O.,S.82).

Man könnte also grundsätzlich beim Vergleich beider Arbeiten an den Einfluß des Bekanntschaftsgrades zwischen Urteilern und Beurteilten als kongruenzsenkende Variable denken, was eher eine realistische Interpretation der IPT nahelegen würde. In eine ähnliche Richtung weisen auch Ergebnisse von HÖHN (1967, S.103). Sie stellte das fast geschlossen negative Bild vom schlechten Schüler (Stereotyp) fest, wir in Ergänzung dazu das geschlossen

positive Bild vom "guten Schüler", (Siehe Abschnitt "Die Stereotypen von Schülern"), was für eindeutige hohe Korrelationen zwischen den verwendeten Eigenschaftsbezeichnungen sprechen müßte, da negative Ausprägungsgrade ausschließlich mit anderen negativen korrelieren, positive Ausprägungsgrade ausschließlich mit anderen positiven.

In aktuellen Urteilen über schlechte Schüler fand sie dagegen eine ganze Reihe von zwar geringen, zum Teil jedoch signifikanten Korrelationen gegen diese "Regel" - positive Ausprägungsgrade eines Merkmals korrelieren positiv mit negativen Ausprägungsgraden eines anderen Merkmals. Besonders auffällig sind in ihren Ergebnissen die gegenläufigen Zusammenhänge zwischen den Eigenschaftsbezeichnungen "faul" und "begabt".
Allein 16,6% ihrer Stichprobe zeigen solche Korrelationen "gegen die Regel", indem sie Schüler als "faul aber begabt", andere als "fleißig aber unbegabt" darstellen, eine Erscheinung, die wir in aktuellen Urteilen ebenfalls festgestellt haben. Würden nun in aktuellen Urteilen ausschließlich semantische Strukturen abgebildet, so müßte man von einem kurzfristigen Bedeutungswandel der Begriffe ausgehen, der nicht sehr wahrscheinlich ist. Wir müssen also in den aktuellen Urteilen wenn schon nicht eine ausschließliche Gestaltung der Urteile durch 'reale' Reize der Beurteilten, so doch wenigstens eine Kontamination semantischer Gemeinsamkeiten der verwendeten Begriffe und 'realer' Reizgegebenheiten annehmen, was für eine realistische Interpretation der IPT spricht. Dabei fassen wir aber den Begriff 'realistisch' nicht in dem Sinn auf, daß 'tatsächliche' Eigenschaften der Reizobjekte die Diskongruenz der Faktorenstrukturen von Stereotypenbeurteilungen und aktuellen Beurteilungen bedingen, sondern

nur einzelne Hinweisreize einzelner aktuell beurteilter Personen.

> Oft, und gerade in den Fällen von Untersuchungen, die Tests zur Überprüfung der Merkmalsvalidität einsetzen, wird gerade auch fachwissenschaftlich jener Begriff "Eigenschaft" benutzt mit allen Implikationen von Konsistenz, Generalität und Universalität, von dem wir im Rahmen der Betrachtung der Alltagsdiagnostik festgestellt haben, daß er hochinferent ist, d.h. besonders anfällig für Urteilstendenzen. Sind aber die fachwissenschaftlichen Feststellungen von "Eigenschaften" qualitativ besser als die des Alltagsdiagnostikers? Objektiver sind sie wohl schon, wobei sich die Objektivität jedoch nur auf die geringe subjektive Beeinflussbarkeit der Messungen und ihrer Auswertung bezieht; dafür sind sie insofern 'subjektiv', als sie sich auf ein bestimmtes theoretisches Konstrukt beziehen, das ebenfalls nur aus bestimmten Verhaltensweisen und -produkten "Eigenschaften" konstruiert wie 'intelligent', 'introvertiert' usw.

Kehren wir wieder zu unserer Vergleichssituation zurück. Die hohe Kongruenz zwischen den Faktoren aus Stereotypenbeurteilungen und aktuellen Beurteilungen relativ 'unbekannter' Schüler hatten wir schon vorausgesehen. Mit der geringsten Bekanntschaft, d.h. mit dem geringen aktuellen Informationsniveau geht in der erzwungenen Urteilssituation eine bestimmte Unsicherheit einher (der Mittelwert der von den Lehrern geschätzten subjektiven Urteilssicherheit betrug 18,2%). In dieser unsicheren Situation werden sich die Lehrer bei der Beurteilung (Markierung von Skalenwerten) auf Ähnlichkeitsschätzungen der vorgegebenen Eigenschaftsbegriffe bezüglich des Reizobjektes "Schüler" allgemein zurückziehen. Diese Ähnlichkeiten lagen auch schon der Beurteilung der Reizobjekte "der, der gute, der schlechte Schüler" zugrunde. Das Auffinden der gleichen Faktoren in beiden Beurteilungen wird zusätzlich noch durch die unsicherheitserzeugte eingeschränkte Skalenbenutzung erleichtert (die durchschnittliche absolute Distanz der Markierungen in aktuellen Urteilen über relativ unbekannte Schüler betrug $M_{|D|} = 1{,}20$, in den Stereotypenbeurteilungen $M_{|D|} = 2{,}44$. Im nächsten Kapitel werden wir uns der Entwicklung der Faktoren mit steigendem Bekanntschaftsgrad zuwenden.

5.5. Kongruenzen der Faktoren aus aktuellen Schülerbeurteilungen mit steigendem Bekanntschaftsgrad

Erwarten wir nun, ausgehend von den Überlegungen im vorigen Kapitel Veränderungen der IPT, so ist es zweckmäßig, nicht nur diese Tatsache festzustellen, sondern auch gemäß unserer Fragestellung die Art dieser Veränderung zu beschreiben.
Wir definieren dazu die IPT nicht mit HAYS (1958) als die Faktorenstruktur der Urteile, da in diesem Fall Ähnlichkeitsrotationen durchgeführt werden müßten sondern mit CLINE (1964) als die konsistente Annahme eines Zusammenhangs mehrerer Persönlichkeitseigenschaften; einschränkend müssten wir hinzufügen, daß diese Eigenschaften zwar von der Gruppe selbst als relevant für Schülerbeurteilungen genannt wurden, daß aber doch nicht jede Eigenschaft für jeden Urteiler gleich relevant ist; wir betrachten die IPT aber mit der Implikation, sie stelle ein Gruppenstereotyp dar. Eine weitere Einschränkung dieser Definition stellt ein gewisser Zwang der Lehrer zur Abgabe expliziter Urteile dar, aus denen wir die IPT rekonstruieren wollen.

Statistisch-methodisch implizieren wir, daß die Faktorenanalyse das adaequate Verfahren zur Darstellung des Sachverhalts ist; besonderes Augenmerk wollen wir auf den Umfang der Varianz richten, die ein Faktor bindet, da wir den Varianzbetrag als Indikator der Wichtigkeit dieses Beurteilungsfaktors für die Urteiler ansehen; in gleichem Sinn wird der gesamte Varianzbeitrag der Faktoren zu einem Merkmal, die Kommunalität dieses Merkmals also, als Maß der Bedeutung des Merkmals für den Urteiler betrachtet.

5.5.1. Die Faktorenlösungen

Auf den folgenden Seiten werden die bereinigten Faktorenmatrizen der Eigenwert- und der 4-Faktoren-Lösung dargestellt, es wurden nur jeweils die Ladungen über .50 und die Prüfgröße $P = a^2/h^2$ nach FÜRNTRATT eingetragen, um die Tabellen übersichtlich zu gestalten. (Originale befinden sich im Anhang).

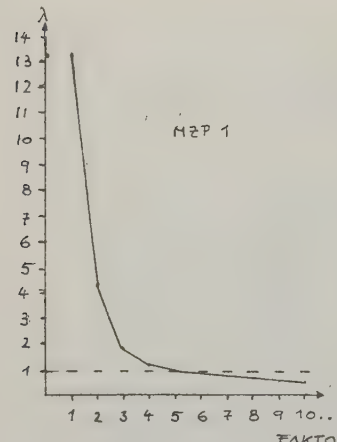
5.5.1.1. Faktorenlösung nach dem Eigenwertkriterium

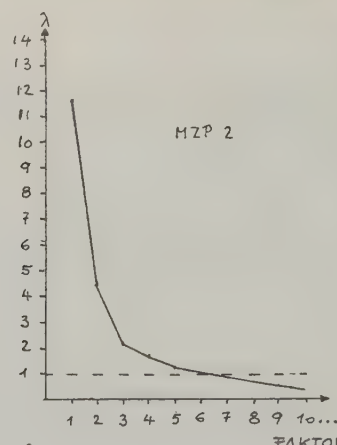
5.5.1.1.1. Faktorenextraktion

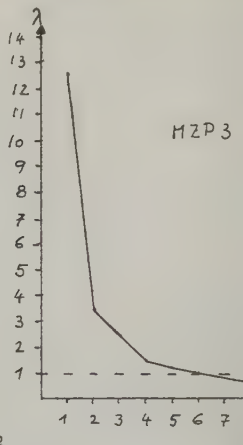
Sehen wir uns die Eigenwertdiagramme oder die Eigenwerttabelle (im Anhang) an, so vermissen wir in allen Diagrammen eine eindeutige Diskontinuität an der Stelle $\lambda = 1$. Ein Ast der Darstellung geht ohne Sprung in den nächsten über, so daß eine Art Hyperbel entsteht, an der eine Rangschätzung nach diesem Kriterium nur schwer möglich ist. Wenn wir allerdings als weiteres Kriterium zur Rangschätzung die Anzahl der Eigenwerte suchen, von der an die Kurve einen linearen Verlauf nimmt, so wird nochmals zu fast allen Meßzeitpunkten die nach den Vorschlägen von FÜRNTRATT (1969) schon im letzten Abschnitt als optimal anzusehende 4-Faktoren-Lösung nahegelegt.

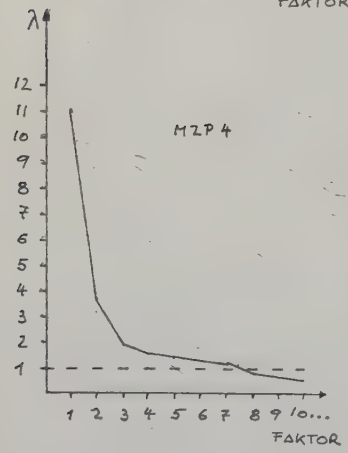
Zunächst wollen wir jedoch zu allen Zeitpunkten die Eigenwertlösung rechnen, da hier durch Extraktionsvorgaben am wenigesten in die Faktorenstruktur eingegriffen wird. Da mindestens zwei oder drei Eigenwerte sich recht nahe am Abbruchkriterium $\lambda = 1$ befinden, könnte es allerdings sein, daß manchmal zuviele Faktoren extrahiert werden, sodaß die letzten Faktoren weniger als drei $a^2/h^2 \geq .50$, aufweisen, womit diese Lösungen zurückgewiesen werden müssen.

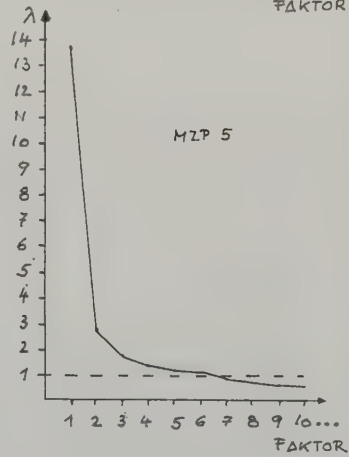
TAB. 9: EIGENWERTDIAGRAMME ZU ALLEN MESSZEITPUNKTEN (MZP

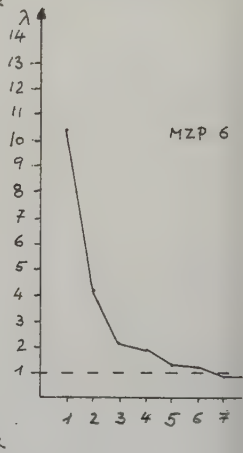


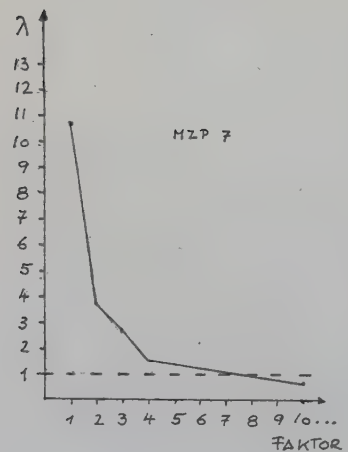


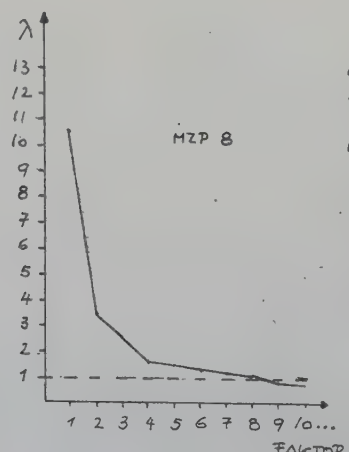


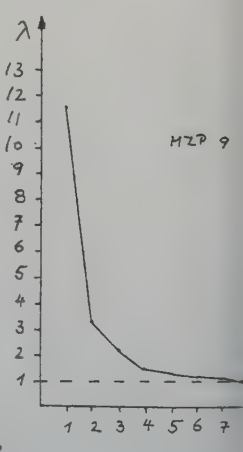




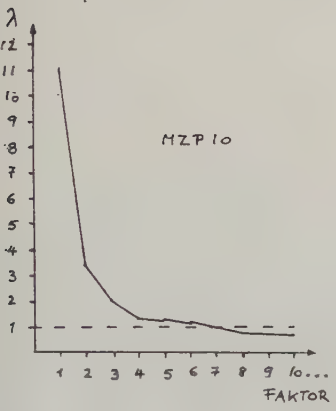


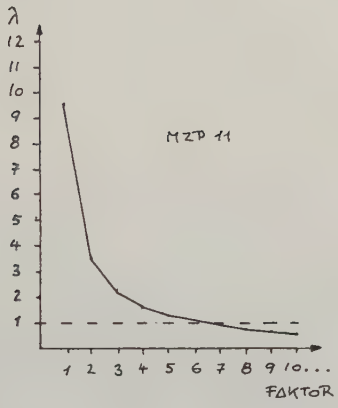






FORTSETZUNG TAB.9





FAKTORENANALYSE MESSZEITPUNKT 1

Kat.	Skala (pos.Pol)	I	II	III	IV	V
Extr.	Varianz %	64,2%	19,7%	7,7%	4,7%	3,7%
1	aufgeweckt	+.76				
2	ehrlich		+ 61		.52	
3	ehrgeizig	+.65				
4	sprachl.begabt					65
5	eig.Mitdenken	+.80				
6	fleißig	.61	+ 66			
7	gt.Lehrerkontakt			+ 76		
8	gt.Mitarbeit	+.68				
9	hübsch			+ 71	.79	
10	höflich		+ 84			
11	intelligent	+.84				
12	kontaktfreudig			+ 75		
13	konzentriert		+ 73			
14	kreativ	+.73				
15	ausdauernd	.60	61			
16	ordentlich		+ 75			
17	paßt s.Klasse an		+ 65			
18	pos.Einst.Schule		+ 70			
19	akzeleriert				+ 68	
20	sensibel					
21	gt.häusl.Milieu		+ 55			
22	unterrichtsstör.		+ 75			
23	unterfordert	+.66			55	
24	begabt	+ 82				
25	selbständig	54			53	
26	sympathisch					55
27	off.Charakter			+ 66		
28	vital				+ 72	
29	zuverlässig		+ 76			

Eigenwertlösung: bereinigte Faktorenmatrix +: $a^2/h^2 \geq .50$
Aus Platzgründen wird in der Folge gerundet und das Komma weggelassen.

Tab.: 10

FAKTORENANALYSE MESSZEITPUNKT 2

Tab.: 11

Kat.	Skala (pos.Pol)	I	II	III	IV	V	VI
Extr.	Varianz %	56,0	20,7	8,5	6,7	4,1	3,8
1	aufgeweckt	+.75					
2	ehrlich		+ 60		.52		
3	ehrgeizig		+ 63				
4	sprachl.begabt						+ 92
5	eig.Mitdenken	+.78					
6	fleißig		+ 77				
7	gt.Lehrerkontakt			+.75			
8	gt.Mitarbeit			.59			
9	hübsch				.79		
10	höflich		+ 73				
11	intelligent	+.84					
12	kontaktfreudig			+.82			
13	konzentriert		+ 77				
14	kreativ	+.82					
15	ausdauernd		+ 73				
16	ordentlich		+ 87				
17	paßt s.Klasse an			.51			
18	pos.Einst.Schule		+ 72				
19	akzeleriert	+.63					
20	sensibel					+ 75	
21	gt.häusl.Milieu						
22	unterrichtsstör.		+ 83				
23	unterfordert	+.83					
24	begabt	+.84					
25	selbständig	+.73					
26	sympathisch				+ 77		
27	off.Charakter			+ 81			
28	vital						
29	zuverlässig		+ 82				

Eigenwertlösung: bereinigte Faktorenmatrix +: $a^2/h^2 \geq .50$
Aus Platzgründen wird in der Folge gerundet und das Komma weggelassen.

FAKTORENANALYSE MESSZEITPUNKT 3

Kat.	Skala (pos.Pol)	I	II	III	IV	V	VI
Extr.	Varianz %	60,6	15,4	10,6	5,2	4,1	4,0
1	aufgeweckt		58	56			
2	ehrlich				63	52	
3	ehrgeizig	+.77					
4	sprachl.begabt						
5	eig.Mitdenken	56	57				
6	fleißig	+ 80					
7	gt.Lehrerkontakt			+ 79			
8	gt.Mitarbeit	56		60			
9	hübsch				+ 53		
10	höflich				+ 56		
11	intelligent		+ 73				
12	kontaktfreudig			+ 86			
13	konzentrierert	+ 80					
14	kreativ		61				
15	ausdauernd	+ 74					
16	ordentlich	+ 79					
17	paßt s.Klasse an				+ 69		
18	pos.Einst.Schule	+ 68					
19	akzeleriert					+ 60	
20	sensibel						+ 82
21	gt.häusl.Milieu					+ 83	
22	unterrichtsstör.	+ 69					
23	unterfordert		+ 75				
24	begabt		+ 74				
25	selbständig		+ 64				
26	sympathisch				+ 79		
27	off.Charakter			+ 79			
28	vital						- 62
29	zuverlässig	+ 85					

Eigenwertlösung: bereinigte Faktorenmatrix +: $a^2/h^2 \geq .50$
Aus Platzgründen wird in der Folge gerundet und das Komma weggelassen.

Tab.: 12

FAKTORENANALYSE MESSZEITPUNKT 4

Kat.	Skala (pos.Pol)	I	II	III	IV	V	VI	VII
Extr.	Varianz %	59,o	15,9	7,6	5,7	4,4	4,3	3,1
1	aufgeweckt		55					
2	ehrlich							
3	ehrgeizig	+ 68						
4	sprachl.begabt							+ 92
5	eig.Mitdenken	57						
6	fleißig	+ 80						
7	gt.Lehrerkontakt		+67					
8	gt. Mitarbeit		+65					
9	hübsch				+ 82			
10	höflich						+ 67	
11	intelligent	54			55			
12	kontaktfreudig		+79					
13	konzentriert	+ 79						
14	kreativ							
15	ausdauernd	+ 78						
16	ordentlich	+ 84						
17	paßt s.Klasse an						+ 72	
18	pos.Einst.Schule	+ 67						
19	akzeleriert			+ 68				
20	sensibel			- 67				
21	gt.häusl.Milieu					+ 95		
22	unterrichtsstör.	+ 65						
23	unterfordert			+ 71				
24	begabt	57						
25	selbständig			61				
26	sympathisch				+ 72			
27	off.Charakter		+78					
28	vital			+ 71				
29	zuverlässig	+ 83						

Eigenwertlösung: bereinigte Faktorenmatrix +: $a^2/h^2 \geq .50$
Aus Platzgründen wird in der Folge gerundet und das Komma weggelassen

FAKTORENANALYSE MESSZEITPUNKT 5

Kat.	Skala (pos.Pol)	I	II	III	IV	V	VI
Extr.	Varianz %	69,8	12,o	7,3	4,3	3,5	3,2
1	aufgeweckt	+ 83					
2	ehrlich	+ 60					
3	ehrgeizig	+ 64		53			
4	sprachl.begabt				+ 63		
5	eig.Mitdenken	+ 88					
6	fleißig	+ 70					
7	gt.Lehrerkontakt		55				
8	gt.Mitarbeit	+ 64					
9	hübsch				+ 81		
10	höflich			+ 58			
11	intelligent	+ 89					
12	kontaktfreudig		+ 84				
13	konzentriert	+ 77					
14	kreativ	+ 78					
15	ausdauernd	+ 87					
16	ordentlich	+ 71					
17	paßt s.Klasse an		54				
18	pos.Einst.Schule	+ 66					
19	akzeleriert	+ 59					
20	sensibel			+ 78			
21	gt.häusl.Milieu					+ 77	
22	unterrichtsstör.						+ 94
23	unterfordert	+ 83					
24	begabt	+ 86					
25	selbständig	+ 89					
26	sympathisch				+ 60		
27	off.Charakter		+ 86				
28	vital		+ 74				
29	zuverlässig	+ 63					

Eigenwertlösung: bereinigte Faktorenmatrix +: $a^2/h2 \geq .50$
Aus Platzgründen wird in der Fole gerundet und das Komma weggelassen

Tab.: 14

FAKTORENANALYSE MESSZEITPUNKT 6

Kat.	Skala (pos.Pol)	I	II	III	IV	V	VI
Extr.	Varianz %	56,1	18,7	8,5	8,3	4,7	3,6
1	aufgeweckt	+ 70					
2	ehrlich					67	
3	ehrgeizig		+ 61				
4	sprachl.begabt				+ 86		
5	eig.Mitdenken	+ 74					
6	fleißig		+ 76				
7	gt.Lehrerkontakt			+ 81			
8	gt.Mitarbeit			+ 63			
9	hübsch					+ 86	
10	höflich						54
11	intelligent	+ 78					
12	kontaktfreudig			+ 86			
13	konzentriert		+ 70				
14	kreativ	+ 79					
15	ausdauernd		+ 65				
16	ordentlich		+ 76				
17	paßt s.Klasse an			+ 60			
18	pos.Einst.Schule		+ 68				
19	akzeleriert	+ 64					
20	sensibel						57
21	gt.häusl.Milieu				+ 67		
22	unterrichtsstör.		+ 75				
23	unterfordert	+ 80					
24	begabt	+ 76					
25	selbständig	+ 85					
26	sympathisch					+ 54	
27	off.Charakter			+ 79			
28	vital	56					
29	zuverlässig		+ 77				

Eigenwertlösung: bereinigte Faktorenmatrix +: $a^2/h^2 \geq .50$
Aus Platzgründen wird in der Folge gerundet und das Komma weggelassen.

FAKTORENANALYSE MESSZEITPUNKT 7

Kat.	Skala (pos.Pol)	I	II	III	IV	V	VI	VII
Extr.	Varianz %	54,7	15,9	8,o	6,o	5,5	5,1	4,9
1	aufgeweckt	+ 66						
2	ehrlich	58		51				
3	ehrgeizig		+ 72					
4	sprachl.begabt						+ 92	
5	eig.Mitdenken	+ 68						
6	fleißig		+ 79					
7	gt.Lehrerkontakt			51				
8	gt.Mitarbeit		+ 59					
9	hübsch			+ 82				
10	höflich							
11	intelligent	+ 82						
12	kontaktfreudig			63				
13	konzentriert		+ 81					
14	kreativ	+ 72						
15	ausdauernd		+ 86					
16	ordentlich		+ 65					
17	paßt s.Klasse an							+ 90
18	pos.Einst.Schule							
19	akzeleriert	+ 71						
20	sensibel				+ 88			
21	gt.häusl.Milieu					+ 69		
22	unterrichtsstör.		560					
23	unterfordert	51						
24	begabt	+ 73						
25	selbständig	+ 78						
26	sympathisch			+ 69				
27	off.Charakter			+ 67				
28	vital							
29	zuverlässig	+ 65						

Eigenwertlösung: bereinigte Faktorenmatrix +: $a^2/h^2 \geq .50$
Aus Platzgründen wird in der Folge gerundet und das Komma weggelassen

Tab.: 16

FAKTORENANALYSE MESSZEITPUNKT 8

Tab.: 17

Kat.	Skala (pos.Pol)	I	II	III	IV	V	VI	VII
Extr.	Varianz %	56,1	14,7	8,2	5,8	5,4	5.o	4,8
1	aufgeweckt	60		58				
2	ehrlich							+ 79
3	ehrgeizig		+ 81					
4	sprachl.begabt							
5	eig.Mitdenken	58						
6	fleißig		+ 71					
7	gt.Lehrerkontakt			+ 74				
8	gt.Mitarbeit							
9	hübsch				+ 68			
10	höflich					+ 76		
11	imtelligent	+ 79						
12	kontaktfreudig			+ 68				
13	konzentriert		+ 73					-
14	kreativ	+ 76						
15	ausdauernd		60					
16	ordentlich		+ 79					
17	paßt s.Klasse an					+ 77		
18	pos.Einst.Schule							
19	akzeleriert	+ 79						
20	sensibel						-83	
21	gt.häusl.Milieu							+ 81
22	unterrichtsstör.		+ 62					
23	unterfordert	+ 62						
24	begabt	58			56			
25	selbständig							
26	sympathisch				+ 76			
27	off.Charakter			+ 80				
28	vital						+ 73	
29	zuverlässig	51						

Eigenwertlösung: bereinigte Faktorenmatrix +: $a^2/h^2 \geq .50$
Aus Platzgründen wird in der Folge gerundet und das Komma weggelassen.

FAKTORENANALYSE MESSZEITPUNKT 9

Tab.: 18

Kat.	Skala (pos.Pol)	I	II	III	IV	V	VI	VII
Extr.	Varianz %	56,6	14,5	8,7	5,4	5,2	4,9	4,7
1	aufgeweckt	+ 64						
2	ehrlich		+ 66					
3	ehrgeizig	+ 60						
4	sprachl.begabt						+ 87	
5	eig.Mitdenken	+ 87						
6	fleißig	58	58					
7	gt.Lehrerkontakt	53		+ 65				
8	gt.Mitarbeit	+ 76						
9	hübsch			+ 83				
10	höflich		+ 70					
11	intelligent	+ 67						
12	kontaktfreudig							
13	konzentriert	+ 74						
14	kreativ	+ 85						
15	ausdauernd	+ 69						
16	ordentlich	+ 64						
17	paßt s.Klasse an							+ 87
18	pos.Einst.Schule		56					
19	akzeleriert					61		
20	sensibel		57					
21	gt.häusl.Milieu		+ 72					
22	unterrichtsstör.		+ 75					
23	unterfordert	. 57						
24	begabt	. 58						
25	selbständig	+ 64						
26	sympathisch			55				
27	off.Charakter				+ 86			
28	vital					+ 79		
29	zuverlässig	59	+ 63					

Eigenwertlösung: bereinigte Faktorenmatrix +: $a^2/h^2 \geq .50$
Aus Platzgründen wird in der Folge gerundet und das Komma weggelassen

FAKTORENANALYSE MESSZEITPUNKT 10

Tab.: 19

Kat.	Skala (pos.Pol)	I	II	III	IV	V	VI
Extr.	Varianz %	58,6	16,5	9,4	6,o	4,9	4,5
1	aufgeweckt	+ 62					
2	ehrlich						
3	ehrgeizig	+ 69					
4	sprachl.begabt				52		54
5	eig.Mitdenken	+ 76					
6	fleißig	54	57				
7	gt.Lehrerkontakt			+ 65			
8	gt.Mitarbeit	+ 71					
9	hübsch				+ 73		
10	höflich		+ 58				
11	intelligent	+ 78					
12	kontaktfreudig			+ 73			
13	konzentriert	53	+ 69				
14	kreativ	+ 81					
15	ausdauernd	+ 73	51				
16	ordentlich		+ 69				
17	paßt s.Klasse an						+ 87
18	pos.Einst.Schule	+ 64					
19	akzeleriert	+ 67					
20	sensibel					+ 75	
21	gt.häusl.Milieu			55			
22	unterrichtsstör.		+ 82				
23	unterfordert	+ 65					
24	begabt	+ 72					
25	selbständig	+ 58					
26	sympathisch				+ 72		
27	off.Charakter			+ 82			
28	vital					- 70	
29	zuverlässig	+ 56					

Eigenwertlösung: bereinigte Faktorenmatrix +: $a^2/h^2 \geq .50$
Aus Platzgründen wird in der Folge gerundet und das Komma weggelassen

FAKTORENANALYSE MESSZEITPUNKT 11

Kat.	Skala (pos.Pol)	I	II	III	IV	V	VI
Extr.	Varianz %	58,2	17,6	9,1	5,4	4,9	4,8
1	aufgeweckt		+ 69				
2	ehrlich			65			
3	ehrgeizig	+ 69					
4	sprachl.begabt						
5	eig.Mitdenken		53				
6	fleißig	+ 74					
7	gt.Lehrerkontakt		+ 77				
8	gt.Mitarbeit	58	58				
9	hübsch			+ 81			
10	höflich						
11	intelligent				51		
12	kontaktfreudig		+ 74				
13	konzentriert	+ 80					
14	kreativ		53				
15	ausdauernd						+ 72
16	ordentlich	+ 80					
17	paßt s.Klasse an						+ 76
18	pos.Einst.Schule	+ 68					
19	akzeleriert				+ 72		
20	sensibel					+ 81	
21	gt.häusl.Milieu			+ 59			
22	unterrichtsstör.	+ 70					
23	unterfordert				+ 59		
24	begabt	55					
25	selbständig		52		54		
26	sympathisch			+ 77			
27	off.Charakter		+ 68				
28	vital					- 74	
29	zuverlössig	51					

Eigenwertlösung: bereinigte Faktorenmatrix +: $a^2/h^2 \geq .50$
Aus Platzgründen wird in der Folge gerundet und das Komma weggelassen.

Zunächst ist zu bemerken, daß die "mathematisch" richtige Faktorenanzahl in den meisten Fällen zu groß war; die Faktoren V, VI und VII klären nur selten einen nennenswerten Teil der Kommunalität ihrer Variablen auf. Das Eigenwertkriterium gibt jedoch nur die obere Grenze der Faktorenanzahl an, sodaß nochmals die Notwendigkeit einer kleineren Lösung betont wird.

5.5.1.1.2. Faktorenidentifikation

Zunächst sollen die Faktoren nach dem "matching index" von CATTELL (1957) identifiziert und geordnet werden. Zwar können schon hier einzelne Beobachtungen gemacht werden, die Identifikation ist jedoch noch recht ungenau, was sich besonders zum Meßzeitpunkt 5 in allen Lösungen bemerkbar macht (vgl. bereinigte Matrizen im vorigen Abschnitt). Wir werden daher zunächst die Faktorkongruenzkoeffizienten berechnen und einige auffällige Punkte besonders hervorheben.

Die Identifikation der Faktoren nach dem "matching index" ergab folgende Rangfolge der Faktoren:

Tab. 21: Faktorenidentifikation nach dem "matching index"

	Messzeitpunkt										
Extraktions-position	1	2	3	4	5	6	7	8	9	10	11
1	I	I	II	(II)	(I)	I	I	I	I	I	(II)
2	II	II	I	III	III	II	II	II	II	II	III
3	III	III	III	(V)	(V)	III	III	III	(III)	III	IV
4	(IV)	(V)	IV	IV	IV	(A)	(V)	(IV)	(B)	(IV)	(C)
5	(V)	(D)	(E)	(E)	(E)	IV	(E)	(F)	(G)	(V)	(V)
6	-	(H)	(V)	(F)	(J)	(K)	(H)	(V)	(H)	(L)	(L)
7	-	-	-	(H)	-	-	(M)	(E)	(M)	-	-

Erläuterung:

(A) - (M): Ein neuer Faktor ist entstanden, zumeist Angliederung einzelner Variablen aus bereits früher bestehenden Faktoren; einander ähnliche Faktoren sind mit dem gleichen Buchstaben gekennzeichnet.

(I) - (V): Ein bereits beschriebener Faktor hat sich durch Ladungen auf andere Variablen stark verändert oder hat eine zu geringe Anzahl wesentlicher Ladungen, sodaß die Identifikation durch den "matching index" unsicher ist.

Die Faktorkongruenzkoeffizienten zwischen dem Meßzeitpunkt 1 und allen folgenden Meßzeitpunkten sowie zwischen aufeinanderfolgenden Meßzeitpunkten zeigen die beiden nächsten Tabellen:

Tab. 22: Kongruenzkoeffizient c zwischen den Meßzeitpunkten

FAKTOR	1/2	1/3	1/4	1/5	1/6	1/7	1/8	1/9	1/10	1/11
I	.94	.90	.85	.96	.95	.93	.86	.94	.97	.82
II	.98	.97	.47	.83	.94	.86	.86	.91	.91	.82
III	.87	.92	.82	.85	.84	.92	.91	.87	.90	.83
IV	.90	.51	.53	.45	.35	.51	.48	.48	.49	.37

Tab. 23: Kongruenz c zwischen den Meßzeitpunkten

FAKTOR	1/2	2/3	3/4	4/5	5/6	6/7	7/8	8/9	9/10	10/11
I	.94	.94	.98	.89	.90	.92	.95	.88	.96	.78
II	.98	.95	.98	.85	.93	.93	.96	.78	.90	.56
III	.87	.93	.94	.89	.94	.75	.85	.76	.82	.85
IV	.90	(.41)	(.80)	(.82)	(.83)	(.77)	(.76)	(.76)	(.75)	(.84)

Zum Meßzeitpunkt 1 wurden nur 5 Faktoren extrahiert, die über die Zeitpunkte hinweg verfolgt wurden. Andere Faktoren traten zu einzelnen Zeitpunkten sporadisch auf, zeigten jedoch keine Stabilität ($c < 45$), weswegen sie zwar in den Tabellen nicht geführt werden können, jedoch in der Diskussion der Meßzeitpunkte behandelt werden. Ebenfalls instabil über die Zeit war Faktor 5, der jedoch schon zum Zeitpunkt 1 als Methodenfaktor aus der weiteren Verfolgung und Interpretation ausgeschieden wurde. Auch Faktor IV zeigte sich in der Form, die er zum Meßzeitpunkt 1 hatte, instabil. Ab Zeitpunkt 2 war jedoch eine gewisse Stabilität dieses neuen Faktors zu beobachten, die in der Diskussion der Meßzeitpunkte besprochen wird.

Tab. 24: Varianzanteile (%) der unrotierten Beurteilungsfaktoren über alle 11 Meßzeitpunkte

	Messzeitpunkte										
FAKTOR	1	2	3	4	5	6	7	8	9	10	11
I	45,9	40,1	43,5	38,3	47,9	35,6	37,5	36,4	39,4	38	33
II	14,8	15,3	11,8	5,4	6,3	13,8	13,0	11,6	11,0	11,8	7,6
III	6,6	7,4	8,9	12,6	9,6	7,1	9,1	8,4	7,1	6,9	12,1
IV	4,4	(6,1)	(5,0)	(5,4)	(4,4)	(4,6)	-	(5,6)	(4,9)	(4,7)	(7,6)

Bei der Durchsicht der Ladungsmatrizen und der Errechnung der Kongruenzkoeffizienten wurde eine Beobachtung gemacht, die Anlaß gibt, die Kongruenzen der Faktoren speziell zu den Zeitpunkten 4,5,10 und 11 auf die Fragestellung 3 hin zu untersuchen (Inhalts- und Bedeutungskonstanz der Faktoren, s.S.54).
Dazu stellen wir die Faktoren in ihrer Gestalt zum Zeitpunkt 1 den noch unidentifizierter Faktoren in ihrer jeweiligen Extraktionsposition zu diesen Zeitpunkten gegenüber und versuchen, eine eindeutige Identifikation durch Kongruenzkoeffizienten vorzunehmen.

MZP 4

		Extraktionsposition 1	2	3	4
MZP 1 Faktor:	I	.85	.69		
	II	.88	.35	.20	.47
	III			.82	
	IV				.53

Kongruenzkoeffizienten der Faktoren

MZP 5

		Extraktionsposition 1	2	3	4
	I	.96	.52		
MZP 1 Faktor:	II	.75	.45	.83	.44
	III		.85	.25	
	IV				.45

M Z P 9

Extraktionsposition

		1	2
MZP 1	I	.93	.64
Faktor	II	.70	.91

M Z P 10

Extraktionsposition

		1	2	3	4
	I	.97	.60		
MZP 1	II	.68	.91		
Faktor	III			.90	
	IV				.49

M Z P 11

Extraktionsposition

		1	2	3	4
	I	.78	.77	.48	.82
MZP 1	II	.82	.39	.56	.41
Faktor	III			.83	
	IV				.37

Der festgestellte hohe Zusammenhang der Faktoren I und II zu den Zeitpunkten 4, 5 und 11 läßt eine eindeutige Identifikation bzw. Trennung beider Faktoren nicht zu oder läßt diese zweifelhaft erscheinen. Von den bedeutsamen Ladungen her bietet sich zu den genannten Zeitpunkten eine andere Interpretation der beiden konfundierten Faktoren an als zum Zeitpunkt 1 für jeden dieser Faktoren getrennt; werden also zu den Zeitpunkten 4, 5 und 11 aktuelle Schülerbeurteilungen erhoben, so muß der Untersucher zu einer anderen Aussage über festgestellte Eigenschaftszusammenhänge in den Urteilen kommen als zu den übrigen Zeitpunkten. Die Dimensionen der Urteile werden doch definiert als "Anzahl der in der Theorie des Lehrers enthaltenen unabhängigen Faktoren" (HOFER, 1969, S.14),was zu den genannten drei Zeitpunkten für die anfänglich festgestellten Faktoren I und II jedoch nicht zutrifft.

Verfolgen wir nun gemäß unserer Fragestellungen das Zustandekommen dieser 'Kontamination' der beiden ersten Faktoren über die Meßzeitpunkte :

Wir unterscheiden in der folgenden Faktorendiskussion zwischen Variablen mit bedeutsamen Ladungen auf einem Faktor ($a > .50$) und Variablen, bei denen ein nennenswerter Teil ihrer Kommunalität durch einen Faktor aufgeklärt wird ($P = a^2/h^2 > .50$), und die nach FÜRNTRATT (1969) vorrangig zur Interpretation des Faktors herangezogen werden sollen.

Meßzeitpunkt 1 : Faktoren I und II sind hinreichend voneinander getrennt, d.h. jeder Faktor erklärt nennenswerte Teile der Kommunalität geladener Variablen auf, wozu der andere Faktor nichts Nennenswertes beiträgt. Faktor I wurd als 'Intelligenz'-Faktor, Faktor II als 'Schuldisziplin'-Faktor identifiziert. Faktor III ist als 'Kontakt'-Faktor zu interpretieren, lädt aber auch bedeutsam die Skala 'hübsch-', was man als emotional-affektive Komponente des Kontaktes deuten könnte.
Faktor IV ist als Faktor der 'psychophysischen Konstitution' mit bedeutsamem P nur auf 'akzeleriert-' und 'vital-' bereits nicht mehr interpretierbar, ebenso Faktor V.

Meßzeitpunkt 2 : Die Skalen 'ehrgeizig-' und 'ausdauernd-' wandern von Faktor I nach II aus. Ebenso verschwindet die Ladung von Faktor I auf 'fleißig-' zugunsten einer höheren Ladung dieser Skala durch Faktor II, der aber schon zum Zeitpunkt 1 einen bedeutsamen Anteil ihrer Kommunalität geklärt hatte. Bedeutsam wird für Faktor I 'akzeleriert-'. Noch immer sind beide Faktoren zu interpretieren wie zum Zeitpunkt 1.
Faktor III ist als der Kontaktfaktor des Mzp 1 zu erkennen, wobei jedoch die Ladung auf 'hübsch-' in einen eigenen Faktor IV auswandert, der zusammen mit der bedeut-

samen Ladung auf 'sympathisch-' als emotional-affektive Komponente der Urteilsreaktion angesehen werden kann (siehe WARR und KNAPPER, 1968, S.2o), jedoch ebenso wie Faktor V nicht die Interpretationskriterien von FÜRNTRATT (1969) erfüllt.

Erstmalig tritt hier ein Faktor VI auf, der jedoch nur die Variable 'sprachlich begabt-' hoch lädt und deren Kommunalität er bedeutsam aufklärt (P=.93). Diese Skala stellt dann auch die zentrale Eigenschaft des Zeitpunktes 2 dar. Man könnte daraus schlußfolgern, daß die Beurteilung etwas differenzierter wird, was wir allerdings einem Testungseffekt zuschreiben; durch die Darbietung der Skala wurde ein Teil der Stichprobe erst auf diese Eigenschaft aufmerksam gemacht, die nun aus Faktor V ausgegliedert wird; nur 65% der Stichprobe hatten diese Eigenschaft als beurteilungsrelevant im Schulbereich genannt.Die Ausgliederung der Skala und Bildung des Faktors VI könnte aber auch die Folge der Extraktion einer zu großen Faktorenzahl sein; die folgende, kleinere Rotationslösung wird zeigen, ob sich dieser Faktor durchsetzt.

Meßzeitpunkt 3: Er ist gekennzeichnet durch die Umkehrung der Bedeutsamkeit der Faktoren I und II; Faktor II 6o,6% der Varianz, Faktor I nur 15,4%. Während Faktor II das originale Ladungsmuster aufweist und als 'Schuldisziplin'-Faktor eindeutig interpretiert werden kann, hat Faktor I nur noch bedeutsame P auf 'intelligent-', 'unterfordert-', 'begabt-' und 'selbständig-'.
Faktor III ist noch als 'Kontakt'-Faktor interpretierbar.
Faktor IV erscheint durch bedeutsame P auf 'hübsch-',

'höflich-', 'paßt sich Klasse an-' und 'sympathisch-' als affektiv betonter 'Anpassungs'-Faktor.
Die Faktoren V und VI sind zu wenig prägnant zur Interpretation, obwohl besonders Faktor VI die gegensätzliche Einschätzung von Sensibilität und Vitalität deutlich macht.

Meßzeitpunkt 4: Hier beginnt die Kontamination der Faktoren I und II. Faktor II 'Schuldisziplin' ist immer noch der bedeutsamere, weist aber bereits bedeutsame Ladungen der Skalen 'eigenes Mitdenken-', 'intelligent' und 'begabt' auf, sodaß seine Interpretation als reiner 'Disziplin'-Faktor bereits schwerfällt.
Als zweitwichtigster Faktor stellt sich der 'Kontakt'-Faktor III heraus, der in seiner ursprünglichen Form erhalten bleibt.
Die nächste Position wird von einem Faktor eingenommen, der mit 'akzeleriert-', 'vital-' und 'unterfordert-' nur noch schwach an den Zeitpunkt 2 erinnert.
Die bedeutsamen P auf den Skalen 'hübsch-' und 'sympathisch-' des nächsten extrahierten Faktors deuten mit ihrem emotional-affektiven Gehalt auf Faktor IV hin, die übrigen Ladungen weisen jedoch nur wenig Kongruenz mit diesem Faktor auf.
Der nächste extrahierte Faktor lädt nur die Skala 'gutes häusliches Milieu-', von der er auch einen wesentlichen Teil ihrer Kommunalität aufklärt. So stellt diese Variable die 'zentrale Eigenschaft' des Zeitpunktes 4 dar.
Mit bedeutsamen P auf 'höflich-' und 'paßt sich der Klasse an-' erinnert der nächste Faktor an den Anpassungs-Faktor, der zum Zeitpunkt 3 als 4. Faktor extrahiert wurde.

Als 7. Faktor wird zu diesem Zeitpunkt ein 'single' mit bedeutsamem P auf 'sprachlich begabt-' ausgegliedert, eine Erscheinung, die wir bereits zum Zeitpunkt 2 als sechsten Faktor beobachtet haben; nun stellt diese Variable nicht mehr die zentrale Eigenschaft dar, obwohl sie mit einem h^2 = .88 nahe an der Kommunalität der zentralen Eigenschaft 'gutes häusliches Milieu-' (h^2 = .93) liegt.

Meßzeitpunkt 5: Die Kontamination der ursprünglichen Faktoren I und II wurde komplettiert. Der erste extrahierte Faktor bindet mit 69,8% die meiste Varianz im Vergleich zu den anderen Zeitpunkten. Das rührt daher, daß er alle Skalen zusammenfaßt, bedeutsam lädt und bedeutende Teile ihrer Kommunalität aufklärt, die anfänglich noch in Faktor I und II sich gegenseitig ausschließend aufgetreten waren. Da er den 'Intelligenz'- und den 'Schuldisziplin'-Faktor vereinigt, müßte man ihn eventuell als 'Eignungs'-Faktor für den Besuch einer höheren Schule auffassen.
Man könnte diesen Faktor für ein methodisches Artefakt halten, mehrere Gründe sprechen jedoch dagegen: Das Extraktionskriterium war das gleiche wie zu den anderen Zeitpunkten; es wurden weder deutlich weniger Faktoren extrahiert (von hier aus könnte man ein 'Zusammenballen' einzelner Faktoren verstehen), noch deutlich mehr Faktoren. Im letzteren Falle wäre es einsehbar gewesen, wenn einzelne Skalen aus bereits bestehenden Faktoren ausgegeliedert worden wären und einen eigenen Faktor gebildet hätten, aber gerade das Gegenteil war der Fall. Wir werden diesen Effekt auch an einer kleineren Rotationslösung prüfen.

Für einen Testungseffekt als Ursache spricht das mit Fortschreiten der Untersuchungshäufigkeit allmählich erfolgende Auswandern der Skalen von einem Faktor zum anderen, gegen einen solchen Testungseffekt spricht die Rückläufigkeit der Kontamination zu den folgenden Meßzeitpunkten und deren rhythmisches Auftreten zu einzelnen Zeitpunkten.
Der zweite Faktor ist der bereits bekannte 'Kontakt-' Faktor III, angereichert durch ein bedeutsames P auf 'vital-' und eine bedeutsame Ladung auf 'paßt sich der Klasse an-', während kein bedeutsamer Teil der Kommunalität von 'guter Lehrerkontakt-' mehr aufgeklärt wird.
Der nächste extrahierte Faktor wurde als der ursprüngliche Faktor II identifiziert, der mit dem dritten Faktor zum Zeitpunkt 5 nur geringfügig kongruenter ist als mit dem 1. Faktor dieses Zeitpunktes.
Er ist jedoch nicht als 'Schuldisziplin'-Faktor zu interpretieren, da er nur zwei bedeutsame P auf 'höflich-' und 'sensibel-' und eine bedeutsame Ladung auf 'ehrgeizig-' enthält. Der vierte extrahierte Faktor klärt bedeutsame Teile der Kommunalität der Variablen 'hübsch-', 'sympathisch-', 'sprachlich begabt-', erinnert damit an den vierten Faktor der Meßzeitpunkte 1, 3 und 4.
Der fünfte extrahierte Faktor lädt auf 'gutes häusliches Milieu-' wie der fünfte Faktor zum Zeitpunkt 4.
Der sechste Faktor ist ebenfalls ein 'single', er lädt nur die Skala 'stört nicht den Unterricht-', weswegen von einer Interpretation abgesehen werden soll.

Meßzeitpunkte 6,7,8: Hier erscheinen die zum Zeitpunkt 5 kontaminierten Faktoren I und II wieder voneinander getrennt und gegeneinander abgesetzt. Der 'Kontakt'-Faktor III wird gut reproduziert, zeigt zum Zeitpunkt 7 jedoch eine starke emotional-affektive Komponente mit bedeutsamen P auf 'hübsch-' und 'sympathisch-'.
Zum Zeitpunkt 6 wurden sechs Faktoren, zu den Zeitpunkten 7 und 8 je sieben Faktoren extrahiert. Interessant wird erst wieder

Meßzeitpunkt 9 : in dem wieder eine Verschmelzung der ursprünglichen Faktoren I und II zu einem 'Eignungs'-Faktor eintritt wie zum Zeitpunkt 5. Der zweite extrahierte Faktor sammelt mit 'stört nicht den Unterricht-', 'zuverlässig-', 'höflich-', 'ehrlich-' Relikte des Faktors II, die jedoch durch ein bedeutsames P auf 'gutes häusliches Milieu-' vor dem sozio-kulturellen Hintergrund gesehen werden, als Einfluß der familiären Erziehung.
Der dritte Faktor ist wieder der 'Kontakt-' Faktor mit emotional-affektiver Beteiligung. Der vierte und die folgenden Faktoren genügen den Interpretationskriterien von FÜRNTRATT (1969) nicht.

Meßzeitpunkt 10 : Faktor I kann hier relativ gut reproduziert und auch als Faktor I interpretiert werden; ein gewisser Einfluß des Faktors II ist mit 'konzentriert-', 'positive Einstellung zur Schule-', 'akzeleriert-' und 'zuverlässig-' jedoch noch zu erkennen. Auch Faktor II läßt sich gut reproduzieren, wie es schon von den Kongruenz-Koeffizienten angedeutet wird.

Der dritte Faktor ist als Faktor III 'Kontakt' zu erkennen, auch hier stimmt die Aussage des Kongruenzkoeffizienten mit der Interpretierbarkeit überein.
'hübsch-' und 'sympathisch-' kennzeichnen den vierten Faktor 'emotional-affektive Bewertung'.
Der fünfte Faktor stellt wieder den Gegensatz zwischen 'sensibel-' und 'vital-' dar, ist aber ebensowenig zu interpretieren wie der sechste Faktor.

Meßzeitpunkt 11 : Die Kongruenzkoeffizienten lassen eine Kontamination der Faktoren I und II vermuten; Faktor I ist jedoch außer durch eine Ladung von a = .55 im ersten Faktor bei dessen Interpretation nicht repräsentiert, taucht aber an vierter Extraktionsposition auf.
Der erste extrahierte Faktor wird als Faktor II identifiziert und ist auch in seiner ursprünglichen Form zu interpretieren.
Der zweite Faktor ist der 'Kontakt-'-Faktor III, während der dritte extrahierte Faktor zum Zeitpunkt 11 wieder den Faktor 'emotional-affektive Bewertung' darstellt.
Bedeutsame Ladungen auf den Skalen 'akzeleriert-', 'unterfordert-', 'selbständig-' und 'intelligent' weisen darauf hin, daß der vierte Faktor den ursprünglichen Faktor I darstellt, der jedoch mit nur noch 5,4% Varianzanteil extrem an Bedeutung verloren hat.
Der fünfte Faktor lädt wie der dritte zum Zeitpunkt 4, und der fünfte zum Zeitpunkt 10 'vital-' und 'sensibel-' gegenläufig auf und stellt nochmals die gegensätzliche Auffassung der Urteiler bezüglich beider Eigenschaften dar; der sechste extrahierte Faktor genügt den Interpretationskriterien nicht.

Ehe wir nun die aufgetretenen Phänomene diskutieren, wollen wir feststellen, ob sie nur durch Besonderheiten der Faktorenextraktion nach dem Eigenwertkriterium entstanden sind; wegen der regelmäßig an Extraktionsposition 5 auftretenden Faktoren, die weniger als drei Skalen aufweisen, deren Kommunalität in bedeutendem Maße von diesen Faktoren aufgeklärt wird, müssen wir vermuten, daß durch das Abbruch-Kriterium $\lambda \geq 1$ zuviele Faktoren extrahiert werden; nach FÜRNTRATT (1969, S.68 f) müssen wir also die 4- und 3-Faktoren-Lösung durchführen.

5.5.1.2. Die 4-Faktoren-Lösung

Für die 4-Faktoren-Lösung gelten - bis auf die jetzt festgelegte Zahl zu extrahierender Faktoren - die gleichen Grundsätze wie schon bei der Faktorenextraktion nach Eigenwertkriterium.

5.5.1.2.1. Faktorenextraktion

Auf den nächsten Seiten werden die bereinigten Faktorenmatrizen der 4-Faktoren-Lösung dargestellt (Originale im Anhang).

Tab. 25: Varianzanteile (%) der unrotierten Beurteilungsfaktoren (4-Faktoren-Lösung)

	Meßzeitpunkte										
	1	2	3	4	5	6	7	8	9	10	11
I	45,9	40,1	43,5	38,3	47,9	35,6	37,5	36,4	39,4	38,0	33,0
II	14,8	15,3	11,8	12,6	9,6	13,8	13,0	11,6	11,0	11,8	12,1
III	6,6	7,4	8,9	6,6	6,3	7,1	9,1	8,4	7,1	6,9	7,6
IV	4,4	6,1	5,0	5,4	4,4	6,7	5,1	5,6	4,9	4,7	5,4
	71,7	68,9	69,2	62,9	68,2	60,2	64,7	62,0	62,4	61,4	58,1

Tab. 26: Relative Varianzanteile (%) der rotierten Beurteilungsfaktoren (4-Faktoren-Lösung)

	1	2	3	4	5	6	7	8	9	10	11
I	67,0	60,9	16,7	19,3	74,4	21,6	19,8	18,0	16,8	66,8	61,7
II	20,5	22,2	65,9	64,7	13,3	59,7	61,5	62,5	68,0	18,5	20,5
III	8,0	9,6	11,6	9,0	7,6	10,0	12,6	11,9	9,7	9,6	11,4
IV	4,6	7,3	5,8	7,0	4,6	8,7	6,1	7,7	5,5	5,3	6,3

FAKTORENANALYSE MESSZEITPUNKT 1

Tab.: 26

Kat.	Skala (pos.Pol)	I		II		III		IV	
Extr. Varianz (%)		67		20,5		8,0		4,6	
Variable		a	p	a	p	a	p	a	p
1	aufgeweckt			71	.68				
2	ehrlich	61	52						
3	ehrgeizig			55	.58				
4	sprachl. begabt								
5	eig.Mitdenken			77	.75				
6	fleißig	70	61	56	.36				
7	gt.Lehrerkontakt					69	78		
8	gt. Mitarbeit			64	.56				
9	hübsch					65	96		
10	höflich	81	94						
11	intelligent			81	.77				
12	kontaktfreudig					76	67		
13	konzentriert	74	70						
14	kreativ			66	68				
15	ausdauernd	66	55	53	39				
16	ordentlich	75	69						
17	paßt s.Klasse an								
18	pos.Einst.Schule	70	60						
19	akzeleriert							65	75
20	sensibel							-50	46
21	gt.häusl.Milieu	54	43						
22	unterrichtsstör.	73	90						
23	unterfordert			66	61				
24	begabt			81	79				
25	selbständig			52	46				
26	sympathisch					55	63		
27	off.Charakter					62	64		
28	vital							62	61
29	zuverlässig	79	77						

4-Faktoren-Lösung
$P = a^2/h^2$

Tab.: 27

Kat. Skala (pos.Pol)	I		II		III		IV	
Extr. Varianz (%)	60,9		22,2		9,6		7,3	
Variable	a	P	a	P	a	P	a	P
1 aufgeweckt			73	65				
2 ehrlich	55	58						
3 ehrgeizig	65	66	42	28				
4 sprachl. begabt								
5 eig.Mitdenken			77	74				
6 fleißig	79	79						
7 gt.Lehrerkontakt					59	51		
8 gt. Mitarbeit			47	38				
9 hübsch							59	93
10 höflich	67	83						
11 intelligent			83	78				
12 kontaktfreudig					90	90		
13 konzentriert	76	69						
14 kreativ			81	88				
15 ausdauernd	74	68	50	31				
16 ordentlich	87	88						
17 paßt s.Klasse an								
18 pos.Einst.Schule	73	80						
19 akzeleriert			62	62				
20 sensibel								
21 gt.häusl.Milieu								
22 unterrichtsstör.	80	91						
23 unterfordert			81	94				
24 begabt			82	78				
25 selbständig			70	78				
26 sympathisch							69	73
27 off.Charakter					78	86		
28 vital					56	58		
29 zuverlässig	80	91						

4-Faktoren-Lösung
$P = a^2/h^2$

Faktorenanalyse Meßzeitpunkt 3

Tab. 28

Kat. Skala (pos.Pol)		I		II		III		IV	
Extr. Varianz (%)		65,9		16,7		11,6		5,8	
Variable		a	P	a	P	a	P	a	P
1	aufgeweckt	64	5ø			57	4o		
2	ehrlich			87	92				
3	ehrgeizig	71	78						
4	sprachl. begabt								
5	eig. Mitdenken	79	76						
6	fleißig	69	64						
7	gt.Lehrerkontakt					8o	82		
8	gt.Mitarbeit	59	53						
9	hübsch								
1o	höflich			62	82				
11	intelligent	84	82						
12	kontaktfreudig					85	91		
13	konzentriert	72	69						
14	kreativ	73	76						
15	ausdauernd	66	62						
16	ordentlich	65	53	57	4o				
17	paßt s.Klasse an			55	72				
18	pos. Einst. Schule	58	49						
19	akzeleriert							43	53
2o	sensibel							64	78
21	gt. häusl.Milieu			51	79				
22	unterrichtsstör.			59	52				
23	unterfordert	65	72						
24	begabt	8o	83						
25	selbständig	6o	48						
26	sympathisch			5o	61				
27	off. Charakter					73	83		
28	vital							69	87
29	zuverlässig	69	57	58	4o				

4 - Faktoren - Lösung

$P = a^2/h^2$

Tab. 29

Kat.Skala (pos.Pol)		I		II		III		IV	
Extr. Varianz (%)		64,7		19,3		9,o		7,o	
Variable		a	P	a	P	a	P	a	P
1	aufgeweckt	57	45			54	4o		
2	ehrlich			56	5o			53	44
3	ehrgeizig								
4	sprachl. begabt								
5	eig.Mitdenken	68	66						
6	fleißig	51	37	65	6o				
7	gt.Lehrerkontakt					65	69		
8	gt.Mitarbeit					55	46		
9	hübsch							66	88
1o	höflich								
11	intelligent	74	67						
12	kontaktfreudig					78	91		
13	konzentriert	72	74						
14	kreativ	52	51						
15	ausdauernd	74	74						
16	ordentlich			68	65				
17	paßt s.Klasse an								
18	pos.Einst.Schule			72	67				
19	akzeleriert	47	65						
2o	sensibel					45	54		
21	gt.häusl.Milieu								
22	unterrichtsstör.			61	75				
23	unterfordert	61	65						
24	begabt	75	73						
25	selbständig	81	79						
26	sympathisch							75	85
27	off.Charakter					79	96		
28	vital					51	56		
29	zuverlässig	61	46	62	48				

4 - Faktoren - Lösung $P = a^2/h^2$

FAKTORENANALYSE MESSZEITPUNKT 5

Kat.	Skala (pos.Pol)	I		II		III		IV	
Extr.Varianz %		74,4		13,3		7,6		4,6	
Variable		a	P	a	P	a	P	a	P
1	aufgeweckt	82	82						
2	ehrlich	58	62						
3	ehrgeizig	66	64						
4	sprachl.begabt					55	91		
5	eig.Mitdenken	89	92						
6	fleißig	73	69						
7	gt.Lehrerkontakt			53	44				
8	gt.Mitarbeit	67	58						
9	hübsch					60	87		
10	höflich	50	51						
11	intelligent	90	95						
12	kontaktfreudig			77	88				
13	konzentriert	80	87						
14	kreativ	75	85						
15	ausdauernd	86	94						
16	ordentlich	75	76						
17	paßt s.Klasse an			47	51				
18	pos.Einst.Schule	73	66						
19	akzeleriert	62	89						
20	sensibel							72	82
21	gt.häusl.Milieu								
22	unterrichtsstör.								
23	unterfordert	86	96						
24	begabt	87	94						
25	selbständig	84	93						
26	sympathisch					57	54		
27	off.Charakter			88	91				
28	vital			62	82				
29	zuverlässig	69	75						

4-Faktoren-Lösung
$P = a^2/h^2$

Tab.: 30

FAKTORENANALYSE MESSZEITPUNKT 6

Kat.	Skala (pos.Pol)	I		II		III		IV	
Extr.	Varianz %	69,7		21,6		10,0		8,7	
Variable		a	P	a	P	a	P	a	P
1	aufgeweckt	65	63						
2	ehrlich			55	76				
3	ehrgeizig			56	46				
4	sprachl.begabt							75	94
5	eig.Mitdenken	70	58						
6	fleißig			74	74				
7	gt.Lehrerkontakt					72	75		
8	gt.Mitarbeit					56	47		
9	hübsch								
10	höflich								
11	intelligent	74	75						
12	kontaktfreudig					89	90		
13	konzentriert			70	71				
14	kreativ	74	76						
15	ausdauernd			66	69				
16	ordentlich			74	89				
17	paßt s.Klasse an					52	75		
18	pos.Einst.Schule			62	65				
19	akzeleriert	57	68						
20	sensibel					-49	58		
21	gt.häusl.Milieu							54	90
22	unterrichtsstör.			61	76				
23	unterfordert	75	83						
24	begabt	75	73						
25	selbständig	82	90						
26	sympathisch								
27	off.Charakter					75	95		
28	vital					64	50		
29	zuverlässig .			73	92				

4-Faktoren-Lösung
$P = a^2/h^2$

Tab.: 31

FATORENANALYSE MESSZEITPUNKT 7

Kat.	Skala (pos.Pol)	I		II		III		IV	
Extr.	Varianz %	61,5		19,8		12,6		6,1	
Variable		a	P	a	P	a	P	a	P
1	aufgeweckt	60	52						
2	ehrlich								
3	ehrgeizig			69	79				
4	sprachl.begabt								
5	eig.Mitdenken	64	64						
6	fleißig			70	70				
7	gt.Lehrerkontakt					61	68		
8	gt.Mitarbeit			57	54				
9	hübsch					64	99		
10	höflich							62	80
11	intelligent	88	82						
12	kontaktfreudig					79	72		
13	konzentriert			76	73				
14	kreativ	73	81						
15	ausdauernd			87	87				
16	ordentlich			55	45				
17	paßt s.Klasse an								
18	pos.Einst.Schule								
19	akzeleriert	61	98						
20	sensibel							73	99
21	gt.häusl.Milieu								
22	unterrichtsstör.							57	54
23	unterfordert	51	69						
24	begabt	75	72						
25	selbständig	70	75						
26	sympathisch					54	72		
27	off.Charakter					82	89		
28	vital							-69	86
29	zuverlässig	59	47						

4-Faktoren-Lösung
$P = a^2/h^2$

FAKTORENANALYSE MESSZEITPUNKT 8

Kat.	Skala (pos.Pol)	I		II		III		IV	
Extr.	Varianz %	62,5		18,0		11,9		7,7	
Variable		a	P	a	P	a	P	a	P
1	aufgeweckt	68	60						
2	ehrlich					53	60		
3	ehrgeizig			69	79				
4	sprachl.begabt								
5	eig.Mitdenken	63	61						
6	fleißig			69	72				
7	gt.Lehrerkontakt					68	78		
8	gt.Mitarbeit	53	51						
9	hübsch					52	58		
10	höflich			59	73				
11	intelligent	90	81						
12	kontaktfreudig					60	59		
13	konzentriert			64	73				
14	kreativ	72	89						
15	ausdauernd			58	53				
16	ordentlich			75	78				
17	paßt s.Klasse an								
18	pos.Einst.Schule								
19	akzeleriert	53	88						
20	sensibel							-63	88
21	gt.häusl.Milieu					63	82		
22	unterrichtsstör.			68	83				
23	unterfordert	71	87						
24	begabt	67	65						
25	selbständig	54	54						
26	sympathisch					54	76		
27	off.Charakter					70	67		
28	vitak							70	89
29	zuverlässig	48	38						

4-Faktoren-Lösung
$P = a^2/h^2$

FAKTORENANALYSE MESSZEITPUNKT 9

Kat.	Skala (pos.Pol)	I		II		III		IV	
Extr.	Varianz %	68		16,8		9,7		5,5	
Variable		a	P	a	P	a	P	a	P
1	aufgeweckt	60	46			61	47		
2	ehrlich			55	46				
3	ehrgeizig	54	52						
4	sprachl.begabt								
5	eig,Mitdenken	77	80						
6	fleißig	59	51	55	44				
7	gt.Lehrerkontakt							51	40
8	gt.Mitarbeit	68	67						
9	hübsch							57	
10	höflich			70	89				
11	intelligent	67	62						
12	kontaktfreudig					79	79		
13	konzentriert	68	76						
14	kreativ	71	91						
15	ausdauernd	79	89						
16	ordentlich	59	55						
17	paßt s.Klasse an								
18	pos.Einst.Schule			54	36			47	75
19	akzeleriert	68	94						
20	sensibel			51	72				
21	gt.häusl.Milieu			61	81				
22	unterrichtsstör.			72	88				
23	unterfordert	59	74						
24	begabt	72	74						
25	selbständig	67	75						
26	sympathisch								
27	off.Charakter					62	98		
28	vital								
29	zuverlässig	53	35	66	56				

4-Faktoren-Lösung
$P = a^2/h^2$

Tab.: 34

FAKTORENANALYSE MESSZEITPUNKT 10

Kat.	Skala (pos.Pol)	I		II		III		IV	
Extr.	Varianz %	66,6		18,5		9,6		5,3	
Variable		a	P	a	P	a	P	a	P
1	aufgeweckt	62	59						
2	ehrlich								
3	ehrgeizig	63	88						
4	sprachl.begabt					65	95		
5	eig.Mitdenken	76	81						
6	fleißig	69	77						
7	gt.Lehrerkontakt			67	75				
8	gt.Mitarbeit	77	84						
9	hübsch			59	74				
10	höflich					49	57		
11	intelligent	75	75						
12	kontaktfreudig			72	77				
13	konzentriert	75	78						
14	kreativ	76	82						
15	ausdauernd	87	95						
16	ordentlich	67	71						
17	paßt s.Klasse an								
18	pos.Einst.Schule	70	71						
19	akzeleriert	56	83						
20	sensibel							71	82
21	gt.häusl.Milieu								
22	unterrichtsstör.								
23	unterfordert	58	69						
24	begabt	70	77						
25	selbständig	51	63						
26	sympathisch			54	54				
27	off.Charakter			56	71				
28	vital							-53	-80
29	zuverlässig	63	71						

4-Faktoren-Lösung
$P = a^2/h^2$

FAKTORENANALYSE MESSZEITPUNKT 11

Kat.	Skala (pos.Pol)	I		II		III		IV	
Extr. Varianz %		61,7		20,5		11,4		6,3	
Variable		a	P	a	P	a	P	a	P
1	aufgeweckt			53					
2	ehrlich			65	71				
3	ehrgeizig	69	95						
4	sprachl.begabt								
5	eig.Mitdenken	61	64						
6	fleißig	82	89						
7	gt.Lehrerkontakt			64	71				
8	gt.Mitarbeit	67	79						
9	hübsch			53	94				
10	höflich					-55	54		
11	intelligent	67	72						
12	kontaktfreudig			63	59				
13	konzentriert	75	94						
14	kreativ	50	47						
15	ausdauernd							50	94
16	ordentlich	71	83						
17	paßt s.Klasse an							53	89
18	pos.Einst.Schule	68	75						
19	akzeleriert								
20	sensibel					-49	87		
21	gt.häusl.Milieu			51	72				
22	unterrichtsstör.					-52	50		
23	unterfordert	65	90						
24	begabt	73	76						
25	selbständig	62	63						
26	sympathisch			75	94				
27	off.Charakter					49	51		
28	vital					65	94		
29	zuverlässig	55	55						

4-Faktoren-Lösung
$P = a^2/h^2$

5.5.1.2.2. Faktorenidentifikation

Die Identifikation nach dem 'matching index' ergab folgendes Bild:

Tab.: 37

Extraktions-position	Meßzeitpunkte 1	2	3	4	5	6	7	8	9	10	11
1	I	I	II	II	I/II	II	II	II	II	I/II	I/II
2	II	II	I	I	III	I	I	I	I	III	III
3	III	III	III	III	A	III	III	III	(III)	B	IV
4	IV	(A)	(IV)	(A)	(IV)	(D)	(IV)	(IV)	(E)	(IV)	(F)

Die nachfolgenden Tabellen zeigen wieder die Faktorkongruenzkoeffizienten der einzelnen Meßzeitpunkte mit Zeitpunkt 1 und der aufeinanderfolgenden Zeitpunkte.

Tab. 38: Kongruenzkoeffizient c zwischen Meßzeitpunkten

Faktor	1 - 2	1 - 3	1 - 4	1 - 5	1 - 6	1 - 7	1 - 8	1 - 9	1 - 10	1 - 11
I	.96	.96	.94	(.81)	.93	,95	.94	.96	(.82)	(.69)
II	.98	.95	.95	(.95)	.94	.87	.95	.94	(.96)	(.62)
III	.85	.80	.76	.71	.78	.80	.81	.75	.71	.58
IV	(.62)	.88	(.77)	.49	(.73)	.70	.78	(.73)	.82	(.86)

Tab. 39: Kongruenzkoeffizient c zwischen Meßzeitpunkten

Faktor	1 - 2	2 - 3	3 - 4	4 - 5	5 - 6	6 - 7	7 - 8	8 - 9	9 - 10	10 - 11
I	.96	.93	.95	(.95)	.88	.93	.95	.93	(.96)	(.92)
II	.98	.91	.90	(.79)	.42	.92	.92	.88	(.51)	(.90)
III	.85	.89	.90	.90	.96	.81	.94	.81	.90	.90
IV	(.62)	(.60)	(.71)	(.58)	(.55)	(.65)	.78	(.84)	.52	(.84)

Erläuterung: Eingeklammerte Koeffizienten zeigen an, daß der Faktor mit dem bisher bekannten Ladungsmuster nicht mehr reproduziert wurde.

In den Kongruenzkoeffizienten zwischen den ursprünglichen Faktoren und den späteren Meßzeitpunkten ist ein fallender Erhaltungsgrad auf die Meßzeitpunkte 5 und 11 hin feststellbar. Verfolgen wir die Entwicklung der ursprünglichen Faktoren anhand der Ladungen in den Faktorenmatrizen, so stellen wir eine ähnliche Entwicklung fest wie bei der Faktorenextraktion nach Eigenwertkriterium:

Faktoren I u.II: (siehe auch bereinigte Matrizen)

Zeitpunkt 1 : ursprüngliches Ladungsmuster

Zeitpunkt 2 : gute Reproduktion des Zeitpunktes 1

Zeitpunkt 3 : hier beginnt die Entwicklung der Verschmelzung der Faktoren I und II. Faktor II wird allmählich zugunsten des Faktors I ausgedünnt, was mit einer Umpolung im Extraktionsrang eingeleitet wird. Hat Faktor I bisher 67, bzw. 60,9 % der extrahierten

Varianz geklärt, so trifft das jetzt nur noch für 16,7 % zu, während Faktor II Vorrang bekommt

Zeitpunkt 4 : Prozeß der Verringerung wesentlicher Ladungen hält an.

Zeitpunkt 5 : Faktor I ist ganz im Faktor II aufgegangen und weist dadurch eine geringere Kongruenz mit Zeitpunkt 1 auf. Im Vergleich mit anderen Zeitpunkten klärt dieser summierte Faktor mit ca.75% die meiste extrahierte Varianz auf.

Zeitpunkt 6 : Hier erscheinen die beiden Faktoren wieder getrennt und recht gut identifizierbar. Die Umpolung im Extraktionsrang, wie sie zum Zeitpunkt 3 das erste Mal beobachtet worden war, hält jedoch an.

Zeitpunkt 7 : Beide Faktoren sind von Zeitpunkt 6 auf 7 gut reproduziert.

Zeitpunkt 8 : Einzelne Skalen von Faktor I wandern nach Faktor II aus ('gute Mitarbeit' (8), 'zuverlässig' (29)

Zeitpunkt 9 : Weitere Vermischung der Faktoren I und II

Zeitpunkt 10: Der Gipfelpunkt der Vermischung beider Faktoren ist wieder erreicht. Der resultierende Gesamtfaktor weist am meisten Ähnlichkeit mit Faktor II auf, weniger mit Faktor I.

Zeitpunkt 11: 1 Jahr nach Meßzeitpunkt 1 weisen die Faktoren nur noch mittlere Kongruenzen mit den anfänglich zum Zeitpunkt 1 extrahierten Faktoren auf. Überraschend hoch ist jedoch zum Zeitpunkt 11 die Reproduktion der Faktoren des Zeitpunktes 10

Die im Vergleich zu den anderen Koeffizienten recht hohen Kongruenzen des Faktors IV sind durch die Akkumulation der Skalen der Faktoren I und II in einem Faktor zu erklären, die zum Zeitpunkt 11 noch anhält.

Faktor III: Er ist der einzige Faktor, der zu jedem Meßzeitpunkt reproduziert wird, wenn auch die Kongruenzen zum ursprünglichen Faktor III. des Zeitpunktes 1 zum Teil erheblich geringer sind als die der ersten beiden Faktoren.

Faktor IV:

Zeitpunkt 1: Ursprünglicher Vergleichsfaktor

Zeitpunkt 2: Teil der Faktoren I und III; an seiner Stelle tritt ein Faktor auf, der die Skalen 'hübsch' und 'sympatisch' lädt, ein Faktor der psychischen und physischen Attraktion, der allerdings nur noch einmal, zum Zeitpunkt 4, auftritt.

Zeitpunkt 3: Der ursprüngliche Faktor ist einigermaßen rein zu beobachten.

Zeitpunkt 4: Ist nach Faktor III ausgewandert, an seiner Stelle wieder der 'Attraktions-Faktor', der dem 'Wärme-Faktor' von ASCH (1946) nahesteht.

Zeitpunkt 5: Skala 'sensibel-' verbleibt im Faktor IV, die anderen Skalen werden Teil von F I und F II.

Zeitpunkt 6: Skala 'akzeleriert-' tritt in Faktor I auf, die Skalen 'sensibel-', 'vital-' sind Teil des 'Kontaktfaktors' III; an seiner Stelle wird ein Faktor extrahiert, der 'sprachlich begabt-' und 'häusl. Milieu-' auflädt, hinter dem etwa eine schichtspezifische Beeinflussung des Urteils stehen könnte.

Zeitpunkt 7: Der ursprüngliche Faktor bleibt weitgehend erhalten, er lädt aber jetzt zusätzlich die Skalen 'höflich-' und 'stört nie den Unterricht-'. Im Gegensatz zum ursprünglichen Faktor ist die Skala 'vital-' negativ geladen, die Skala 'sensibel-' positiv, was in Kombination mit den beiden neu hinzugekommenen Skalen einen Faktor 'Zurückhaltung vs. Aufdringlichkeit' vermuten läßt.

Zeitpunkt 8: Der ursprüngliche Faktor taucht wieder auf, ist aber nur mit einer Kongruenz von .78 reproduziert, die Skala 'akzeleriert-' fehlt.

Zeitpunkt 9: Skala 'akzeleriert-' ist in I, Skala 'sensibel-' in II aufgegangen, 'vital-' lädt nirgends bedeutsam. Stattdessen wird ein Faktor extrahiert, der mit Faktor III zu .73 kongruent ist, dessen Ladungen nur geringe Teile der Kommunalität der geladenen Variablen aufklären. Unter Berücksichtigung der obengenannten Interpretationskriterien und -auflagen soll von einer Benennung dieses Faktor abgesehen werden.

Zeitpunkt 10: Der ursprüngliche Faktor IV ist wieder reproduziert, Skala 'akzeleriert-' tritt jedoch in I auf.

Zeitpunkt 11: Faktor IV ist nach Faktor III ausgewandert und bildet dort einen Faktor, der von den geladenen Variablen her große Ähnlichkeit hat mit dem zum Zeitpunkt 7 auf dem 4. Extraktionsrang stehenden Faktor 'Zurückhaltung vs. Aufdringlichkeit'. Statt dessen wird mit den Skalen 'ausdauernd-' und 'paßt sich Klasse an-' ein Faktor extrahiert, der auf die Zuschreibung einer 'langfristigen vs. kurzfristigen Einordnung' sowohl in die Arbeitsanforderungen als auch in den sozialen Kontext schließen läßt.

Aufgrund der engen Zusammenhänge zwischen den Faktoren I und II speziell zum Zeitpunkt 5 muß eine nichtorthogonale Struktur vermutet werden.
Da eine Oblimax-Rotation nicht verfügbar war, wurden zum Zeitpunkt 5 schiefwinklige Rotationen mit verschiedenen Winkelstellungen der Hauptachsen vorgenommen.
Die geringsten Faktorinterkorrelationen erbrachte die Rotation mit $\measuredangle$ = -.0,5 (cos 120°):

		FAKTOR 2	3	4
FAKTOR	1	.17	.03	.27
	2	-	-.06	.21
	3		-	.16

(Originalmatrix im Anhang).

Von den Interpretationen her konnten jedoch die Faktoren I und II trotzdem nicht getrennt werden, der 'Eignungs'-Faktor blieb bestehen.

5.5.1.2.3. Zusammenfassung der 4-Faktoren-Lösung

Wie in der Lösung nach dem Eigenwert-Kriterium ist auch bei der 4-Faktoren-Lösung eine Signifikanz und Reproduktion von nur 3 Faktoren feststellbar; - Faktor I: 'Schuldisziplin', Faktor II: 'Geistige Regsamkeit', Faktor III: 'Zuwendung'. Der vierte extrahierte Faktor wechselt inhaltlich über die Zeitpunkte hinweg so erheblich, daß von einer Konsistenz nicht gesprochen werden kann.
Insgesamt sind zu verschiedenen Zeitpunkten Akkumulationen geladenen Skalen speziell der ursprünglichen Faktoren I und zu beobachten. Diese Anhäufung, die für eine etwas geringer werdende Differenziertheit der Zuschreibung von Variablen i den aktuellen Beurteilungen spricht, kann wohl kaum mit ein sich steigernden Untersuchungsmüdigkeit der Vpn erklärt wer da keine kontinuierliche Steigerung über den Untersuchungsz raum zu beobachten ist. Als weiterer Beleg gegen diese Beha tung mag die hohe Kongruenz der Faktoren am Ende der Beurte lungsserie (10) und in der Einzelmessung nach einem halben Jahr (11) dienen, denn nachdem die Erhebung der Urteile 6 Monate ausgesetzt hatte, kann ja wohl von Untersuchungsmüdi keit aus Gründen der zusätzlichen Arbeitsbelastung nicht me gesprochen werden. Gegen eine steigende Sorglosigkeit der Beurteiler als Grund für die angesprochene Akkumulation spricht auch die in allen drei fraglichen Meßzeitpunkten (5,10,11) typische Verschmelzung der Faktoren I und II, die sich speziell zwischen den Zeitpunkten 10 und 11 in recht hohen Kongruenzen als eine überdauernde Erscheinung manifestiert.

.2.3.1. Die zentrale Dimension

Als 'zentrale Dimension' soll analog zu HOFER (S.49) diejenige angesehen werden, auf die der größte Varianzanteil entfällt und die demzufolge die wichtigste Dimension für den Lehrer bei seinen Beurteilungen darstellt.
In der Korrespondenzuntersuchung wurde aus den Korrelationen der aktuellen Schülerbeurteilungen als zentrale Dimension der Faktor "Disziplinviertheit vs. Undisziplinviertheit" mit Aufklärung von 31 % der extrahierten Varianz errechnet. Zu den Zeitpunkten 1 und 2 ist dieser Faktor als "Schuldisziplin" in der vorliegenden Arbeit ebenfalls feststellbar. Diese Übereinstimmung basiert keineswegs nur auf einer mehr oder weniger zufälligen Benennungsgleichheit, sondern auf der Tatsache, daß dieser Faktor in den beiden verglichenen Untersuchungen bei verschiedenen Personenstichproben weitgehend die gleichen Skalen lädt, nämlich "fleißig-", "zuverlässig-", "ordentlich-", "konzentriert-", "aufmerksam-", "ehrgeizig-". In dieser Untersuchung werden außerdem noch die Skalen "höflich-", "ausdauernd-", "stört selten den Unterricht-", geladen, was diesen Faktor in die Nähe des in der Literatur aufgeführten Faktors "Willenskontrolle" (PAWLIK, 1968, S.398) bringt, der die Eigenschaften Verantwortlichkeit, Ausdauer, Gründlichkeit, Verläßlichkeit umfaßt und der auch von CATTEL u. COAN (1957), CATTELL u. GRUEN (1953) und PETERSON u. CATTEL (siehe HOFER, 1969, S.61) gefunden wurde. Allerdings stellt dieser Faktor, wie aus den Faktorenmatrizen zu ersehen ist, nur eben an diesen beiden Meßzeitpunkten 1 und 2 die zentrale Dimension dar. Zu den Zeitpunkten 3,4,6,7,8 steht der Faktor "Geistige Regsamkeit" im Vordergrund mit 'intelligent-', 'begabt-', 'einfallsreich-', (hier etwa 'kreativ') als einigen der geladenen Skalen, die in beiden Untersuchungen auftauchen.

Auf das Uberwiegen dieses Faktors zu den meisten Zeitpunkten wird in der Interpretation der Ergebnisse noch gesondert eingegangen.
Zu einigen Zeitpunkten (5,9,10,11) wiederum akkumulierten beide bis dahin abwechselnd zentralen Dimensionen zu einem summativen Faktor I/II, der alle Aspekte von Geistiger Regsamkeit und Schuldisziplin vereinigte und oben schon als 'Eignung für eine höhere Schulbildung' angesprochen wurde. Auch dieses Phänomen wird im Kapitel 'Interpretation' eine genauere Betrachtung erfahren müssen.

Zusammenfassend ist festzustellen, daß über die Meßzeitpunkte hinweg keine durchgehende zentrale Dimension im Sinne HOFERs in dieser Stichprobe zu beobachten war, daß sich vielmehr zwei Dimensionen ablösten und zu manchen Meßzeitpunkten sogar zu einer einzigen, umfassenden Dimension zusammenschlossen.

.2:3.2. Die zentrale Eigenschaft

Bei HOFER (S.50) setzt sich mit 'uninteressiert-interessiert' eine zentrale Eigenschaft, definiert als die Skala mit der höchsten Kommunalität, gut von den anderen Eigenschaften ab. Da im vorliegenden Zahlenmaterial keine so eindeutigen Differenzen zwischen den Kommunalitäten vorliegen, soll eine Häufigkeitsauszählung der vier höchsten Kommunalitäten pro Zeitpunkt wenigstens einen Hinweis geben. Die Häufigkeitsrangfolge lautet: 'intelligent-', 'begabt-', 'kontaktfreudig-', 'zuverlässig-','fleißig-', erinnert damit stark an die anfangs aufgeführte 'reflektierte' Rangreihe' auf Frage 2(S.69), was trotz der ständigen Veränderung der Kommunalitätenhöhe noch einmal die augenscheinliche Wichtigkeit dieser Eigenschaften für die beurteilenden Lehrer unterstreicht.

5.6. Einzeluntersuchungen über weitere Kennwerte der IPT

Untersuchungen, die sich mit dem Wirken von IPTn in pädagogischen Urteilen befassen, äußern sich stets besorgt über die dadurch eintretende Einschränkung der Abbildung der freien Variation der Schülermerkmale, die die Gültigkeit der Urteile stark herabsetzt. Abgesehen davon, daß Beurteilungs-'fehler', wie sie aus der Anwendung der IPT in aktuellen Urteilen resultieren, eigentlich erst durch den Einsatz von hochinferenten Eigenschaftsbegriffen in Untersuchungen provoziert werden, scheint diese Kritik von der Modellvorstellung auszugehen, daß die im Lehrerurteil korrelierten Merkmale 'in Wirklichkeit' in der Schülerpersönlichkeit nicht korreliert sind. Der Beweis für die Berechtigung dieser Vorstellung kann jedoch nicht erbracht werden, da es sich beim Urteiler in jedem Fall um Interpretationen von Beobachtungen handelt, von denen nur wenige aus dem physischen und sozio-kulturellen Bereich einer objektiven Betrachtung oder gar Messung zugänglich sind.

Man wird sich also mit Abbildungen der 'wirklichen Persönlichkeit' begnügen müssen und die Persönlichkeit grundsätzlich unter dem Aspekt definieren, wie sie beurteilt wird. Unter diesem Aspekt ist die Persönlichkeit des Beurteilten auch relevant für den Entwurf von Handlungsplänen durch den Urteiler.

Erkennt man nun, daß der Urteilende bezüglich der wahrgenommenen Persönlichkeit Vorannahmen aufweist, die sein 'Bild vom Anderen' weit stärker bestimmen als objektive Informationen, daß diese Informationen nun den Anstoß geben für einen im Urteiler ablaufenden Prozeß, verschiebt sich die Frage nach der 'wirklichen Persönlichkeit' der Beurteilten auf die Frage, wie bestimmte Merkmale des Urteilers auf das Urteil einwirken.

Wie wir Modellen der Personwahrnehmung entnehmen, wirken im Urteilsprozeß explizite und implizite Persönlichkeitskonzepte, Bedürfnisse, Einstellungen und Vorurteile des Urteilers, die person-, rollen- oder gesellschaftsspezifisch ausgerichtet sind. In der Folge wird nun versucht werden, über die impliziten Persönlichkeitskonzepte hinaus weitere Faktoren (nicht statistisch-methodisch aufgefaßt) und ihre Beziehungen zu Urteilen festzustellen.

Zunächst wären hier, der Theorie von CRONBACH (1955) folgend, die erwartungserzeugenden 'Schülerstereotypen' zu untersuchen.

5.6.1. Die Stereotypen

Mit den im ersten Teil der Untersuchung genannten Fragen 4, 5 und 6 sollte versucht werden, die stereotypen Einstellungen der Lehrer bezüglich 'typischer' Schüler festzustellen. Während die Fragen nach "dem guten " und "dem schlechten" Schüler ohne Vorbehalte von den Lehrern beantwortet wurden, erklärten sich sechs jüngere Lehrer (jünger als 33 Jahre) nicht imstande, "den Schüler" der 1. Klasse des Gymnasiums mit jeweils einem Rating-Skalen-Punkt zu charakterisieren, da die Persönlichkeiten zu verschieden seien.
Diese Lehrer wurden in späteren Vergleichen als Sondergruppe geführt; für sie wurde auch die Zusatz-Instruktion entworfen (s.S.74), nach der sie sich eher bereiterklärten, diese Frage zu beantworten. Der Begründung dieser anfänglichen Weigerung nach war allerdings weniger die tatsächliche Beobachtung der Vielfalt der Ausprägungen von Persönlichkeitsmerkmalen und die Unmöglichkeit der Reduktion die Ursache für diese Besorgnis, als eher der Argwohn, sie sollten zu vorschnellen Generalisierungen provoziert werden, um ihnen danach mangelnde Differenziertheit bei der Sicht ihrer Schüler vorzuhalten. Die Vorbehalte resultierten nach Angaben der Lehrer aus der Kenntnis psychologischer Persönlichkeitstheorien.

1.1. Stereotyp: "der schlechte Schüler" der 1. Gymnasialklasse

Die Mediane der Angaben auf diese Fragen bewegten sich, wie zu erwarten war, vorwiegend im negativen Bereich, jedoch waren folgende Besonderheiten zu bemerken.

1) Die Meinung der Lehrer über den schlechten Schüler ist nicht durchweg negativ. Besonders hervorgehoben werden die positiven Ausprägungen in den Skalen: hübsch, offener Charakter, vital, kontaktfreudig, kreativ und akzeleriert.

2) Zweigipfelige Verteilungen liegen in den Skalen höflich, begabt, fleißig, intelligent vor:

(Abb.6) Verteilung der Lehrerurteile über "den schlechten Schüler"

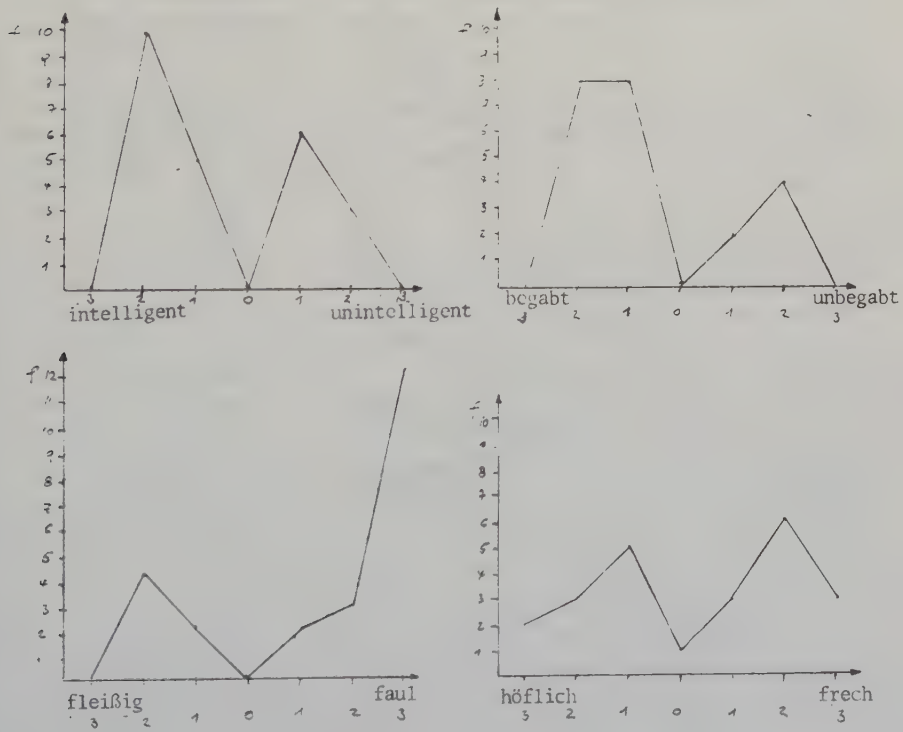
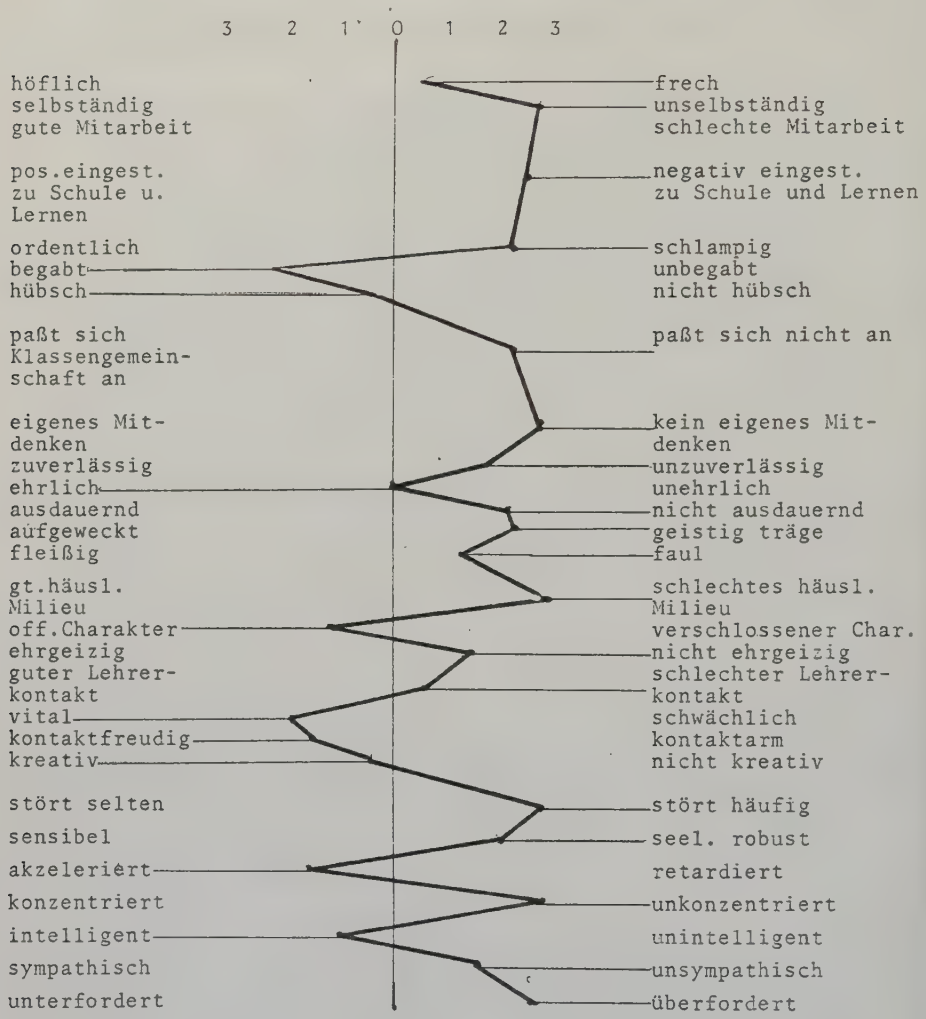


Abb.7 : Stereotyp: "der schlechte Schüler"

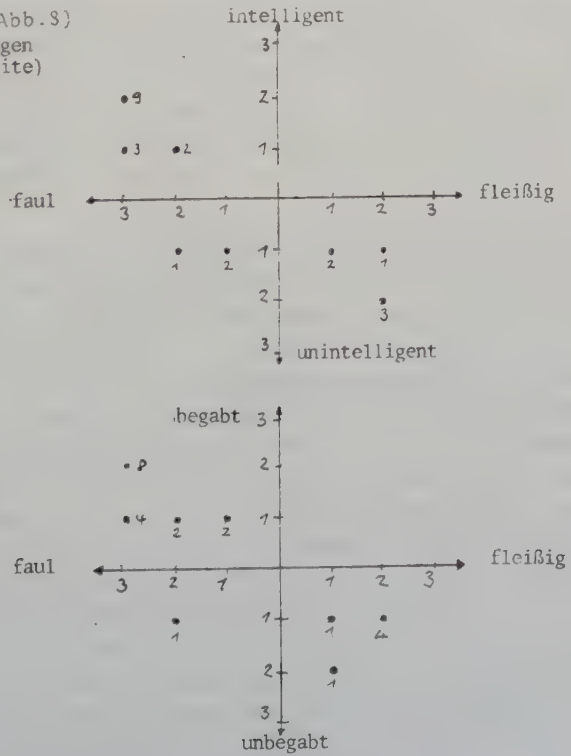


Es scheint also, daß sich - im Gegensatz zu den anderen Merkmalen - bei diesen vier Merkmalen die Meinungen der Lehrer bezüglich des schlechten Schülers unterscheiden. Erstaunlich ist die Verteilung der Werte auf den Skalen 'intelligent-' und 'begabt-'. Es wird anscheinend selbst dem schlechten Gymnasiasten des 1. Schuljahres von der Mehrheit der Lehrer eine gewisse Intelligenz und Begabung zugesprochen. Bedenkt man das Ausleseverfahren für höhere Schulen, nach dem nur Schüler vorgeschlagen werden, die einen bestimmten Notendurchschnitt erreicht haben, also schon einmal bewiesen haben, daß sie etwas leisten können, so wird diese Ansicht der Lehrer verständlich.

Ein Test der beobachteten dichotomisierten Häufigkeiten gegen die genau entgegengesetzten Erwartungen im gleichen Verhältnis erbrachte ein signifikantes Abweichen von dieser Erwartung in den Skalen 'intelligent-' und 'begabt-', nicht jedoch in den Skalen 'fleißig-' und 'höflich-'.

Betrachten wir die Zusammenhänge zwischen Intelligenz und Begabung einerseits und der Anstrengung des Schülers (fleißig-faul) andererseits, so bietet sich folgendes Bild: (Abb.8)

(Erläuterungen nächste Seite)



Erläuterung: Die eingetragenen Punkte geben die Kombinationen (x;y) der Achsenvariablen in den 23 einzelnen Lehrerurteilen über den "schlechten Schüler" an. Die Zahlen neben den Punkten geben die Häufigkeiten der N = 23 Lehrerurteile an, in denen diese Kombination (x;y) auftrat.

Die Häufigkeiten in den Quadranten unterscheiden sich zum Teil ganz erheblich; die dichotomisierten Markierungen auf den Skalen 'intelligent - unintelligent' und 'fleißig - faul' korrelieren zu $r_\varphi = -.64$, 'begabt - unbegabt' mit 'fleißig - faul' zu $r_\varphi = -.67$. Die χ^2 sind in beiden Fällen signifikant, was andeutet, daß die Verteilung der beobachteten Häufigkeiten bei der Beurteilung des 'schlechten Schülers' überzufällig ist. Man kann also mit einiger Vorsicht (wegen des geringen Stichprobenumfanges) sagen, daß bei den Lehrern die Tendenz zu zwei Auffassungen des schlechten Schülers besteht: die häufigste ist die des 'intelligenten faulen' schlechten Schülers, die von 61 % der Stichprobe vertreten wird, für 'unintelligent und fleißig' halten 26 % der Stichprobe den schlechten Schüler, während 13 % ihn als 'unintelligent und faul' sehen. Ähnliches gilt für den 'begabten faulen' und den 'unbegabten fleißigen' schlechten Schüler, was allerdings höchstens für einen synonymen Gebrauch der Begriffe 'intelligent' und 'begabt' spricht.

Die Beobachtung dieser typischen Merkmalsausprägungen, die zur Erklärung des geringen Schulerfolges herangezogen werden, deckt sich mit Beobachtungen von HÖHN (1967,S.103), die diese Erklärungsversuche ebenfalls schildert, und von HECKHAUSEN (1972, S.967), der bei Schulversagen die Tendenz der Lehrer zu externaler Kausalattribution aufgezeigt hat: am häufigsten wird den Schülern die Verantwortung für Mißerfolg zugeschrieben; geringe Intelligenz oder Begabung als Grund anzugeben, mag für einen Volksschullehrer möglich sein, ein Gymnasiallehrer hat es jedoch mit Schülern zu tun, die, wie wir gesehen haben, bezüglich dieser Merkmale schon eine positive Auslese darstellen, womit die Begründung "mangelnde Fähigkeiten" nicht so leicht mehr heranzuziehen ist. "Besonders auffällig ist, welch geringe Rolle sie (die Gymnasiallehrer; Zusatz d. Verfassers) dem Begabungsmangel zumessen, während er für die

Volksschullehrer das prägnante Merkmal des schlechten Schülers ist, über das ... große Einigkeit besteht ..." (HÖHN, 1967, S. 102).

Vielmehr wird der Mißerfolg durch eine besonders geringe Anstrengungsbereitschaft des Schülers zu erklären versucht. Das Attribuierungsschema, das KLEBER (1978, S.607) anführt, nennt als weiteren Bereich, mit dessen Hilfe der Lehrer mangelnden Schulerfolg erklärt, "außerschulische Lernbedingungen des Schülers", das Elternhaus also. Untersuchen wir den Stereotyp des 'schlechten Schülers' unserer Stichprobe daraufhin, dann ist zu bemerken, daß neben dem geringen Mittelwert der Skalen, die Anstrengungsbereitschaft andeuten (fleißig, Mitarbeit), einer der niedrigsten Mittelwerte in der Skala 'häusliches Milieu' festzustellen ist. Ferner sei daran erinnert, daß in der Faktorenanalyse der Stereotypen ein eigener Faktor errechnet wurde, der nur diese eine Skala repräsentierte.

5.6.1.2. Stereotyp: "der gute Schüler" der 1. Gymnasialklasse

Am Stereotyp des 'guten Schülers' sind keinerlei negative Merkmalsausprägungen festzustellen, er wird durchgehend positiv geschildert - kein Wunder, denn er stellt ja für den Lehrer ein Erfolgserlebnis dar, der ihn auch signifikant sympathischer findet als den schlechten Schüler.
Im Gegensatz zu den Typen des schlechten Schülers ist er begabt, intelligent und fleißig und sein häusliches Milieu ist signifikant besser als bei jenem. Darüber hinaus fallen sein eigenes Mitdenken, seine Konzentration und seine Ausdauer auf, in denen er sich ebenfalls bedeutsam vom schlechten Schüler unterscheidet (Vorzeichentest).
Niedere Ausprägungen werden ihm nur in den Merkmalen 'vital', 'ehrlich' (wahrscheinlich irrelevant für seine Charakterisierung) und in Merkmalen zugeschrieben, die sich auf den Sozialkontakt beziehen: Klassenanpassung, offener Charakter, kontaktfreudig, guter Lehrerkontakt. Er wird als etwas ruhiger, zurückgezogener und sensibler als der schlechte Schüler gesehen, allerdings ohne sich in diesen Merkmalen signifikant von ihm zu unterscheiden.
Bemerkenswert ist hier, daß er sich nach Ansicht der Lehrer in den Merkmalen 'intelligent' und 'begabt' nicht bedeutsam vom schlechten Schüler unterscheidet, wohl aber in seinem Arbeitsverhalten. Hier kann wohl die Aussage gewagt werden, daß nach Meinung der Lehrer weniger die Intelligenz und Begabung über den Schulerfolg entscheidet als das Arbeitsverhalten. Diese Aussage kann von den Ergebnissen wenigstens der ersten vier Wochen der Untersuchung gestützt werden. Interpretiert man die Extraktions-Position eines Faktors als dessen Wichtigkeit, so ist wenigstens im genannten Zeitraum der Faktor 'Schuldisziplin' der wichtigste in den Augen der Lehrer.

Abb.9: Stereotyp: "der gute Schüler"

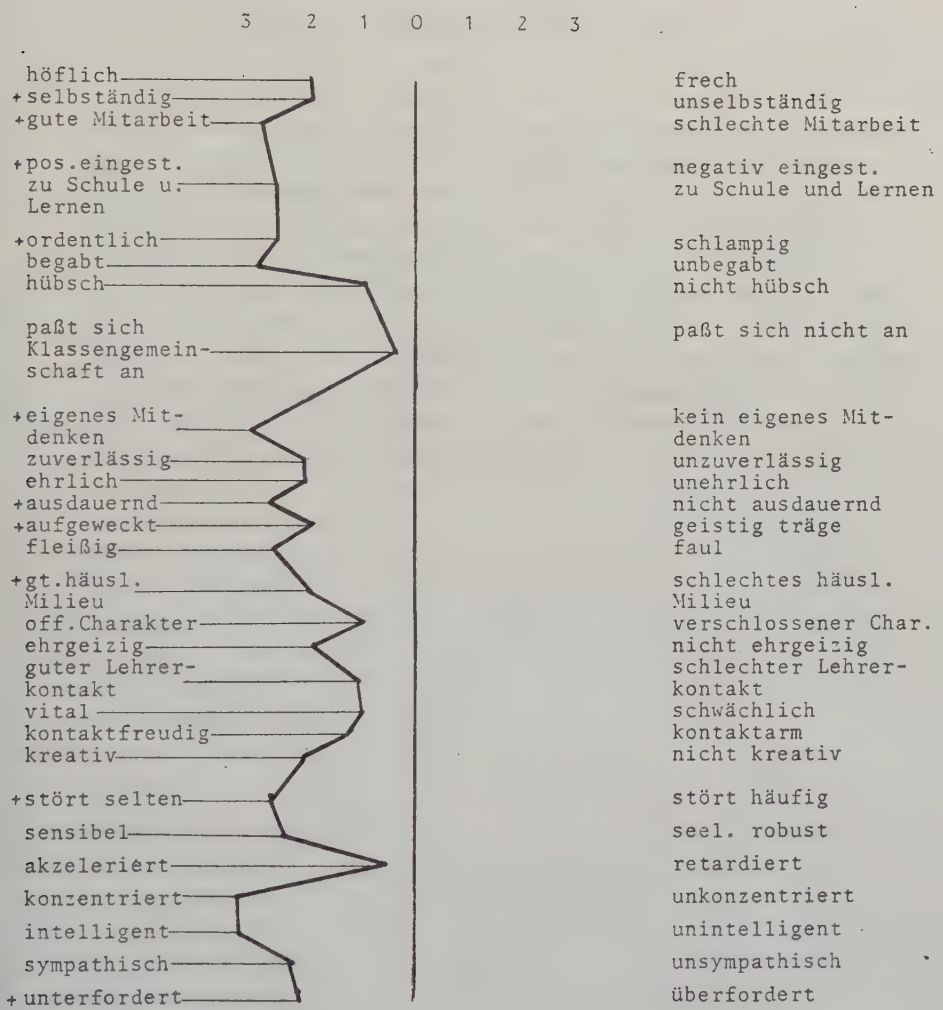


Abb.9 : Stereotyp "der gute Schüler"; Medianwerte
N = 23;
+: Signifikanter Unterschied zum "schlechten Schüler" (Vorzeichentest)

Zu späteren Zeitpunkten (ab 3. Meßzeitpunkt) dominiert zwar der Faktor 'geistige Regsamkeit' in den aktuellen Lehrerurteilen, wird jedoch von Skalen des immer wichtiger werdenden Faktors 'Schuldisziplin' unterwandert und vermischt sich, wie oben dargestellt, zu einzelnen zeitlichen Kulminationspunkten (etwa Meßzeitpunkte 5 und 10) vollständig mit diesen.
Von dieser Entwicklung her kann also keine eindeutige Aussage über die höhere Wichtigkeit eines der beiden Faktoren gemacht werden, wie sie vielleicht der Vergleich der Stereotypen anzeigen würde, vielmehr kann auf diesen beiden Wegen zum Ausdruck kommen, daß die Lehrer der Meinung sind, daß - grob vereinfacht ausgedrückt - Intelligenz und Fleiß nötig sind, um eine gute Schulleistung zu gewährleisten und daß beide Faktoren sich ergänzend aufeinander bezogen sind.

1.3. Stereotyp: "der Schüler" der 1. Gymnasialklasse

Grundsätzlich fällt an diesem Stereotyp auf, daß es stark negativ gefärbt ist. Der Mittelwert über das Profil hinweg beträgt $\bar{x} = -0,22$; deutlicher wird das negative Bild, wenn man die positiven und negativen Abweichungen vom Mittelpunkt der Skalen einander gegenüberstellt: Die Summe der Abweichungen in positiver Richtung beträgt $D_+=11,5$, $N=12$, die in negativer Richtung $D_-=17,6$, $N=15$. Positiv vermerkt wird erstaunlicherweise 'höflich' - was wohl vor dem Hintergrund der älteren Gymnasiasten zu sehen ist-, besondere Beachtung finden auch 'kontaktfreudig' und 'vital' - bei der letzteren Eigenschaft ist jedoch fraglich, ob sie als positiv angesehen wird; ziehen wir die Faktorenanalysen zu Rate, so sehen wir, daß 'vital' meist in Gesellschaft von 'Kontakt'-Variablen auftritt und 'sensibel' ausschließt. (Siehe Faktoren III und IV der 4-Faktoren-Lösung S. 174 ff) 'Vital' scheint also eher die Umschreibung einer aus ungezügelter Energie herrührenden "Umtriebigkeit" zu sein, wie es ein Lehrer in einem Gespräch formulierte -, eine Eigenschaft also, die in einer auch auf 'Schuldisziplin' ausgerichteten Institution nicht positiv bewertet wird.

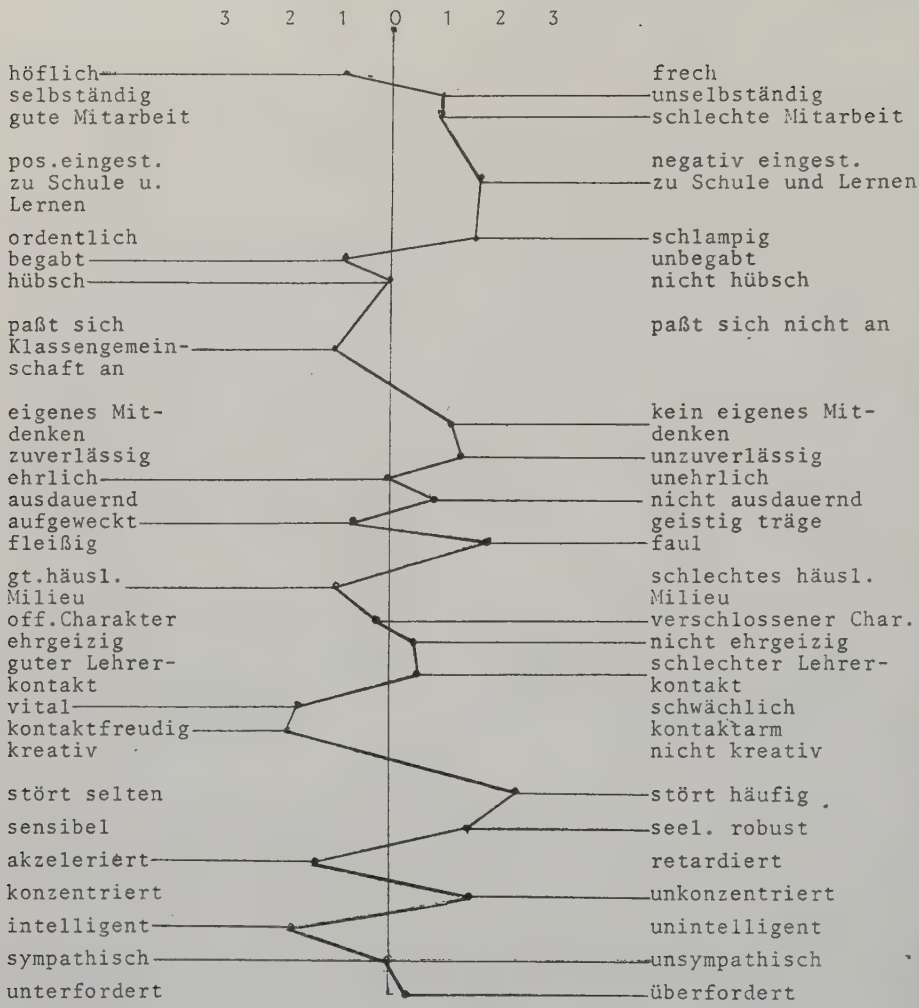


Abb.1o : Stereotyp "der Schüler" der 1. Gymnasialklasse (Medianwerte)

Eine positive Meinung der Lehrer besteht jedoch bezüglich der Intelligenz und Begabung 'des' Schülers der ersten Gymnasialklasse, eine Beobachtung, die wir schon beim 'schlechten Schüler' zu erklären versucht haben.

Besonders negativ wird 'der' Schüler bezüglich seiner Arbeits- und Schuldisziplin eingeschätzt: die Variablen 'stört häufig', 'faul', 'unzuverlässig', 'schlampig', 'unkonzentriert' weisen die höchsten negativen Ausprägungsgrade auf.

'Der' Schüler wird also in seiner geistigen Regsamkeit durchaus positiv gesehen, insgesamt jedoch resultiert aus Gründen der geringen Einschätzung seiner Arbeitsdisziplin ein pessimistischer Gesamteindruck.

5.6.2. Stereotypen und Lehrermerkmale

Es läßt sich denken, daß eine gewisse Affinität bestehen könnte zwischen Persönlichkeitsmerkmalen von Lehrern und der subjektiven Wichtigkeit von Schülermerkmalen für diese Lehrer. Man könnte sich z.B. vorstellen, daß für besonders korrekte Lehrer ein wesentlicher Grund für Schulerfolg und -versagen von Schülern in deren größerer oder geringerer Zuverlässigkeit zu suchen ist oder daß emotional besonders stabile Lehrer den Ausprägungsgrad der emotionalen Stabilität vs Labilität eher zur Erklärung von Erfolg oder Versagen von Schülern heranziehen.
Um solche Appetenzen oder Aversionen festzustellen, wurden die Ratings der Lehrer auf den Sterotypenskalen einem Extremgruppenvergleich unterzogen.

Zur Herstellung der Extremgruppen dienten

a) die T-Werte der Lehrer auf den Skalen des FPI: die Lehrer-Stichprobe wurde nach ihren T-Werten in jeder Skala gedrittelt und die scores der Stereotypenskalen der Extremgruppe bis zum 33. Centil (low-scores) wurden den scores der Extremgruppe über dem 66. Centil (high-scores) mit dem U-Test von Mann-Whitney gegenübergestellt. Der U-Test mit Rangaufteilung wurde wegen der z.T. starken Asymmetrie der Ratings auf den Stereotypenskalen gewählt, die in manchen Fällen einen Betrag von S = 1,47 erreicht.

b) das Alter der Lehrer: die beiden Extremgruppen wurden ebenfalls als das obere und das untere Drittel der Lehrerstichprobe, geordnet nach Alter, definiert; das weitere Verfahren wurde wie unter a) durchgeführt.

c) die Geschlechtszugehörigkeit der Lehrer, wobei nach Maßgabe der Stichprobe-Zusammensetzung, männlich:weiblich = 69,57 : 30,43 die Extremgruppen im Gegensatz zu a) und b) unterschiedlichen Umfang aufwiesen.

In allen 1218 Rechnungen, die aus dieser Anordnung resultierten, zeigten sich jedoch bezüglich der in den Stereotypenskalen angekreuzten Ausprägungsgrade von Schülereigenschaften keine signifikanten Unterschiede zwischen den jeweiligen Extremgruppen, weswegen die Ergebnisse hier auch nicht dargestellt werden sollen. Auch einzelne Rechen-Stichproben mit den veränderten Extremgruppendefinitionen C25/C75 (oberes und unteres Viertel der Stichprobe) und C20/C80 (oberes und unteres Fünftel der Stichprobe) brachten keine signifikanten Unterschiede zwischen den neu definierten Extremgruppen.
Als Resultat kann man also zusammenfassen, daß keine systematischen Zusammenhänge festzustellen waren zwischen Persönlichkeitsmerkmalen der Lehrer und Persönlichkeitsmerkmalen, die diese Lehrer den Schüler-Stereotypen 'der gute Schüler', 'der schlechte Schüler', 'der Schüler' der ersten Gymnasialklasse zuschrieben.

Dieses Ergebnis mag zum Teil untersuchungsspezifisch bedingt sein (geringer Stichprobenumfang, also geringer Extremgruppenumfang mit - bezüglich Zufallsabweichungen - empfindlichen Kennwerten; für signifikante Unterschiede nicht ausreichende cluster-Bildung in beiden Variablen-Gruppen; durch Reduktion der Maßzahlen auf Ordinalskalen-Niveau bedingter Informationsverlust etc.); andererseits unterstreicht es nochmals die zu Anfang dieser Studie gemachte Aussage, daß es sich bei den hier untersuchten Stereotypen eher um rollenspezifische Stereotypen handelt als um persönlichkeitsspezifische, weswegen ihrer Untersuchung anhand von Persönlichkeitsmerkmalen der Lehrer kein Erfolg beschieden war. Geht man davon aus, daß rollenspezifische Ansichten die Beurteilung der fiktiven Schüler-Typen stärker beeinflussen als spezielle Persönlichkeitsmerkmale der Lehrer, so müssen wir einer ähnlich geplanten Untersuchung wie der obigen, diesmal aber nicht auf der Basis von Stereotypen sondern von aktuellen Beurteilungen realer Schüler ebenso schlechte Erfolgschancen einräumen. Am einflußreichsten könnte hier wohl die Lehrer-Variable 'Dienstalter' sein, wenn man von der Annahme ausgeht, daß die Lehrer mit geringem Dienstalter die rollenspezifischen Verhaltensvorschriften, Meinungen,

Ansichten noch nicht so sehr internalisiert haben wie die dienstälteren, die ihre Rolle bereits gelernt haben. Ein ähnliches Ergebnis, das sich in dieser Richtung bewegt, scheint sich bei HOFER (a.a.o., S.54) anzudeuten, der feststellte, daß Lehrer, die neben anderen Merkmalen auch das Merkmal 'jung' aufwiesen (wobei in der Regel die Dienstjahre mit dem Lebensalter hoch korrelieren), eine differenziertere Sichtweise ihrer Schüler zu haben scheinen und eher auch Merkmale des Sozialverhaltens im Urteil über ihre Schüler berücksichtigen.

Da dieser Nachweis dem genannten Autor aber trotz großen rechnerischen Aufwandes bei einer wesentlich größeren Stichprobe als der in dieser Studie vorliegenden nicht eindeutig gelang, soll später ein solcher Versuch nur an einer eng spezifizierten Teilstichprobe mit den Daten der aktuellen Beurteilungen erfolgen (S.235).

Mit diesem Kapitel soll die Ausschöpfung der Informationen über Persönlichkeitsmerkmale der Lehrer und der Versuch abgeschlossen werden, Beziehungen zwischen diesen Informationen und den Daten herzustellen, aus denen die IPT extrahiert wurde.

In den folgenden Kapiteln sollen die Beurteilungen anhand ihrer aus den Rohwerten errechneten Kennziffern (Korrelationen, Mittelwerte, Streuungen) auf Tendenzen hin untersucht werden, die auf faktorenanalytischem Wege nicht zu erkennen waren, da es sich hierbei ja nur um einen Vergleich der Urteilsstrukturen zwischen den Zeitpunkten handelte. Zuletzt sollen noch die Beurteilungen auffälliger Schüler auf Besonderheiten in ihrer Entwicklung über die Meßzeitpunkte hin untersucht werden. Eine generelle Analyse der Beurteilungsentwicklungen wird nicht beabsichtigt, da die Verfolgung der 13.340 Merkmals-Verlaufsgestalten (Merkmalsentwicklungen über die 11 Meßzeitpunkte hinweg) einmal den thematische Vorsatz der Arbeit nicht erfüllt, zum anderen als bloße Ergänzung der eigentlichen Fragestellung den Rahmen der Arbeit sprengen würde.

5.7. Einfluß der subjektiven Urteilssicherheit

Weiter oben war schon darauf hingewiesen worden, daß die Urteilssicherheit der Urteiler sich in den abgegebenen Urteilen dergestalt niederschlagen könnte, daß die Urteile im Zustand der Unsicherheit eine geringere Abweichung vom Indifferenz-Punkt der Skala (O-Punkt) zeigen werden als im Zustand einer höheren subjektiven Sicherheit des Urteilers bezüglich der 'Richtigkeit' des abgegebenen Urteils.
Die Urteilssicherheit war für jeden Beurteilten nur einmal für die ganze Beurteilungsliste festgehalten und vom Urteiler in einer Skala von O-100 % eingetragen worden. Der angegebene Sicherheitswert wurde mit der durchschnittlichen absoluten Abweichung der in den Beurteilungslisten eingetragenen Werte vom O-Punkt korreliert.

Übersicht 1: Subjektive Urteilssicherheit der Lehrer

Meßzeitpunkte		1	2	3	4	5	6	7	8	9	10	11
Sicherheit (%);	M =	18,2	18,6	19,2	24,0	20,4	37,5	43,7	56,3	42,1	45,5	78,4
|Diff.| Scores-Nullpunkt	M =	1,20	1,01	1,14	1,23	0,9	1,42	1,78	1,95	1,82	1,90	2,04

N = 460

Die Korrelation ergab einen Wert von .906 s.01 mit einem Konfidenzintervall von 0.53 - .98.
Das bedeutet, daß zwischen der subjektiven Urteilssicherheit und der Tendenz zur Mitte der Skala in dieser Stichprobe ein enger Zusammenhang besteht. Je unsicherer der Urteiler, desto näher hält er sich in seinen Urteilen am Indifferenzpunkt auf, je sicherer, desto 'ausladender' werden seine Urteile. Das ist an sich kein aufsehenerregendes Ergebnis, sollte aber zum Anlaß genommen werden, zu überlegen, ob nicht die IPT von Urteilern bezüglich ihnen weniger gut bekannten Personen etwa dadurch zustandekommt, daß die effektiv ausgenutzte Skalenbreite wesentlich geringer ist als bei bekannteren Personen, und damit die Korrelationswahrscheinlichkeit sich erhöht.

Leider finden sich in diesbezüglichen Untersuchungen nur Angaben über die Korrelationen der Skalen untereinander, nicht aber Angaben über Mittelwerte und Streuungen.

Erstaunlich sind hier einige Werte, die die Sicherheitsschätzung betreffen. Speziell der erste Wert ist schon recht hoch, obwohl doch eigentlich fast unbekannte Personen vorliegen. Das könnte die Annahme unterstützen, daß die Anwendung der IPT, die ja für diesen Zeitpunkt nachgewiesen ist, ein subjektives Sicherheitsgefühl bezüglich des abgegebenen Urteils erzeugt und entspricht etwa der Auffassung der IPT als Ordnungs- und Strukturierungshilfe in unstrukturierten Situationen. Bemerkenswert ist ferner der Sicherheitswert zum Zeitpunkt 10 und 11. Man müßte eigentlich annehmen, daß zu diesen Zeitpunkten, an denen es um recht folgenschwere Entscheidungen für die Beurteilten geht, ein höheres Gefühl der Sicherheit mit dem Persönlichkeitsurteil verbunden ist.
Die Lehrer-Vpn sind hier aber noch überwiegend unsicher, bzw. unsicher geworden, denn zum Zeitpunkt 8 waren sie schon sicherer als zum Zeitpunkt 10. Vergleichen wir die Bewegung der Sicherheitswerte auf die Zeitpunkte 5 und 10 hin mit der Bewegung der Faktorenstrukturen, so stellen wir eine gewisse Parallelität fest, denn zu den genannten Zeitpunkten akkumulieren sich die Faktoren I und II, wie wir es bei jeder Faktorenlösung festgestellt haben.
Suchen wir nach parallelen Erscheinungen im Schulalltag, so stellen wir gerade zu den Zeitpunkten 5, 10 und 11 Entscheidungssituationen fest: zum Zeitpunkt 5 werden die "blauen Briefe" verschickt, die den Eltern der Kinder ankündigen, daß das Bestehen des Probehalbjahres gefährdet ist, wenn ihre Kinder nicht bessere Leistungen zeigen als bisher, zu den Zeitpunkten 10 und 11 standen Zeugnisformulierungen ins Haus. Dieses Zusammentreffen der genannten Entscheidungen könnte hier dahingehend interpretiert werden, daß ein erzwungenes starkes Reflektieren über die Schüler zu einer größeren Unsicherheit bei der Anwendung der gewohnten IPT führt, was sich im relativen Sinken der subjektiven Urteilssicherheit und in einer Globalisierung der Urteilsstruktur manifestiert.
Diese Interpretation der Werte mag aber etwas zu spekulativ sein,

wenn man bedenkt, daß einmal dieses Reflektieren nur einige Schüler betrifft, die Gefahr laufen, die Schule verlassen zu müssen, zum anderen die extrahierte IPT ja eine Essenz aus vielen IPTn über die einzelnen Schüler ist, die je nach einzelnen Schülermerkmalen stärker oder weniger stark differieren.

5.8. Modifikation der Beurteilung über den Meßzeitraum

Im folgenden Kapitel werden wir uns mit den Beurteilungen der Lehrer auf den einzelnen Skalen und deren Veränderung befassen. Da jetzt die Rohwerte betrachtet werden, sind einige statistische Vorbemerkungen notwendig.

Grundsätzlich kann angenommen werden, daß die Meßwerte in einer Zeitreihe nicht unabhängig voneinander sind. Das war im faktoriellen Teil der Untersuchung noch nicht so sehr wichtig, da ja die Faktoren innerhalb jeden Zeitpunktes festgestellt wurden und dann nur die Ladungsmuster zwischen den Meßzeitpunkten verglichen wurden.

Jetzt aber haben wir es mit den Rohwerten aufeinanderfolgender Messungen zu tun, und hier ist die Abhängigkeit nicht nur eine angenommene, sondern ist durch den Betrag der Korrelation nachweisbar. Die allgemeine Richtung der Veränderung zwischen den Meßzeitpunkten ist allerdings dadurch nicht angegeben; diese Richtungsänderung versuchen wir durch die Verfolgung der Mittelwerte zu erfassen. Manche Vpn werden nach unten vom vorherigen Meßwert abweichen, manche nach oben hin, durchschnittlich aber braucht sich über alle Vpn hinweg der Mittelwert nicht zu verändern.
Wenn er sich aber doch signifikant verändert, so kann das entweder in einer Bedeutungsänderung der verwendeten Variablen (Eigenschaftsbegriffe) oder in neuen Informationen über den Beurteilten begründet sein.

Es ist kaum anzunehmen, daß sich der semantische Gehalt der Begriffe innerhalb einer Frist von 14 Tagen zwischen den Meßzeitpunkten verändert; in den Faktorenanalysen haben wir ja wenigstens die wichtigsten Faktoren zu jedem Meßzeitpunkt mit ausreichender Kongruenz wiedergefunden. Was sich also verändern könnte, ist die Bedeutung der verwendeten Eigenschaften im Sinne ihrer Wichtigkeit für die gerade aktuelle Beurteilung. So mag es in der Entscheidungssituation, ob ein Schüler für den Besuch

des Gymnasiums vorgeschlagen wird, oder ob er weiter im Gymnasium bleiben soll, weniger wichtig für den Lehrer sein, die Kontaktfreude oder Kreativität des Schülers als für diese Fragestellung weniger relevante Eigenschaften zu beurteilen, als eher solche Eigenschaften, die von zentraler Bedeutung für diese Entscheidung sind - wie z.B. das Arbeitsverhalten oder die Intelligenz.
Die Umgruppierung der festgestellten Faktoren und der erhöhte Betrag an Varianzerklärung speziell durch den ersten Faktor würde für diese Annahme sprechen.
Ferner ist denkbar, daß in solchen Entscheidungssituationen die Urteile nicht nur auf einer Perzeption, sondern auf einer Apperzeption der Beurteilten beruhen, daß also Indikatoren für Eigenschaften gezielt beobachtet und verwertet werden, was eine weitere Zentrierung der Urteile auf dem Lehrer relevant erscheinende Eigenschaften vermuten läßt.

Von Interesse könnte hier sein, welche Variablen im Meßzeitraum zuerst Änderungen zeigen, welche Änderungen an den"kritischen" Punkten 4/5 (explizite Zwischenbeurteilung) und 10/11 (endgültige Entscheidung über den Verbleib der Schüler am Gymnasium) eintreten, ob es Variablen gibt, in denen sich die Urteile überhaupt nicht signifikant verändern.

Um die Veränderung des durchschnittlichen Lehrerurteils festzustellen, sollte eigentlich der von HELMREICH (1977,S.52) dargestellte t-Test für Meßwertdifferenzen abhängiger Stichproben gerechnet werden; die Differenzen speziell zwischen den Zeitpunkten 1 - 2 - 3 erwiesen sich jedoch als nicht normalverteilt, weswegen der WILCOXON-Test gerechnet wurde. (Die aus beiden Tests resultierenden Signifikanzen unterschieden sich jedoch nur geringfügig). Da wir an der Veränderung der Anfangsbilder der Lehrer von den Schülern interessiert sind, wird jeder Zeitpunkt mit dem Zeitpunkt 1 verglichen. Von einem Vergleich der

Kennzahlen versprechen wir uns kein prägnantes Ergebnis, da sich die Mittelwerte tendenziell gleichgerichtet, aber nicht signifikant von Zeitpunkt zu Zeitpunkt verändern und somit ein Gleichbleiben des Bildes signalisieren können; unsere Fragestellung bezieht sich aber auf die Veränderung des Bildes vom Schüler mit steigendem Bekanntschaftsgrad bei relativ weitgehender Kongruenz der Urteilsfaktoren, wie wir festgestellt haben.
Wenn wir in der Folge die Lehrer als Gruppe betrachten, so geschieht das im Einklang mit dem Untersuchungsgang des faktorenanalytischen Teils der Arbeit. Sicher könnten aus der differenzierten Bearbeitung der IPT-Kennwerte, wie sie CRONBACH (1955) auffasst, noch viele wertvolle Informationen entnommen werden; eine differenziertere Bearbeitung würde aber den Rahmen der Arbeit sprengen, da diese eigentlich ausschließlich faktorenanalytisch angelegt war.

Es müsste nun angenommen werden, daß der Mittelwert der Urteile zu einem Zeitpunkt gleich O ist. Die Abweichungen vom Mittelpunkt der Skala stellen die 'Strenge' ('error of negative leniency'; GUILFORD,1954) bzw. die 'Milde' ('generosity error'; CRONBACH, 1960) dar. Sehen wir uns die Tabelle auf der nächsten Seite an, dann sehen wir, daß die Schüler mit einem zum Teil erheblichen Bonus bereits zum Zeitpunkt 1 beurteilt werden, Gründe dafür könnten die individuellen Bezugspunkte der einzelnen Urteiler als Funktion aktueller und gespeicherter Informationen, die Anforderungsnormen, die eigene Einschätzung des Beurteilers bezüglich des Merkmals und vor allem die verbale Verankerung der Skala darstellen. Während sich die letzten beiden beeinflussenden Faktoren mit großer Wahrscheinlichkeit während des Untersuchungszeitraumes nicht oder nur geringfügig verändern werden, wird der größte Teil der Veränderung - über alle Lehrer gesehen - wohl auf Einflüsse aktueller und gespeicherter Informationen zurückzuführen sein.

Zum Zeitpunkt 1, zu dem die Lehrer praktisch keine Informationen über die Schüler hatten, fällt die durchschnittliche positive Abweichung von 0,637 Skalenpunkten vom Nullpunkt auf, ein Indiz für eine relative Milde der Beurteilung, evtl. aus der zu diesem Zeitpunkt herrschenden Unsicherheit heraus.

Variable	MGES.:			Sig. 1-2		Sig. 1-3		Sig. 1-4		Sig. 1-5		Sig. 1-6		Sig. 1-7		Sig. 1-8		Sig. 1-9		Sig. 1-10		Sig. 1-11	M Ges.
		Messzeitpunkte																					
		,637		,637		,637		,675	.01	,597	.01	,616	.01	,659	.01	,694	.01	,667	.01	,744	.01	,542	.01	
aufgeweckt	1	,176		,68	.05	,74		,74		,59		,56		,73		,82		,61		,55		,79	.05	,688
ehrlich	2	,196		,199		1,09		1,06	.05	1,05		1,10		1,27		1,53		1,26	.05	1,27		,91	.01	1,13
ehrgeizig	3	,147		,167		,155	•	,166		,165	.05	,167		,163	.01	,183	.01	,166		,142	.05	,137	.05	,598
sprachl. begabt	4	,151		,120		,107		,132	.01	,125		,138		,124		,116		,130		,130	.05	,137	.01	,263
eig. Mitdenken	5	,152		,155		,153		,155		,157		,148		,158	.01	,146		,151	.01	,136		,156		,515
fleissig	6	,179		,178		,177		,187	.05	,182		,181		,175		,197		,178		,169	.01	,148	.01	,173
gt. Lehrerkontakt	7	,113		,167		,177		,182		,170	.05	,193		,172		,174	.01	,167	.01	,169	.01	,168		,738
gt. Mitarbeit	8	,153		,156		,158		,166	.05	,146		,155		,147	.01	,157	.01	,159		,162	.05	,144		,548
hübsch	9	,176		,197	.01	,189		,185		,177	.05	,180		1,00		,196		,179		,182		,175		1851
höflich	10	,185		,195	.05	,196		1,06	.01	1,08		1,16		1,09		1,10	.01	1,24		1,11		,178	.01	1,034
intelligent	11	,174		,172		,166		,167		,167		,171		,181	.01	,181	.05	,183		,163	.01	,190		,736
kontaktfreudig	12	,167		,161	.01	,174	.01	,164		,175		,158		,172		,162		,162	.01	,153	.05	,168		,651
konzentriert	13	,140		,127		,147	.05	,147		,141		,143		,128		,126		,136		,140		,100	.01	,375
kreativ	14	,120		,131		,138		,124		,136		,119		,127		,126		,127		,128	.05	,146	.01	,293
ausdauernd	15	,131		,136		,141	.01	,158	.01	,143	.01	,155	.01	,135		,149		,142		,145	.01	,108	.01	412
ordentlich	16	,142		,183		,184		,189		,177		,180		,195		,197		,181	.05	,178	.01	,158		,831
paßt s. Klasse an	17	,180		,179		,192	.01	,193	.05	,178	.01	,186		,143		1,05	.01	,170	.01	,158	.01	,155		,781
pos. Einst. Schule	18	1,09		1,16		1,02		1,06		,196		,197		1,24		1,30		1,17		1,17		,194	.01	,998
akzeleriert	19	,130		1,04		,119		,120		,114	.05	,112	.05	,121		,113		,113		,109		,117		,165
sensibel	20	,134		,127	.01	,125		,114		-,107	.05	-,123		-,102	.01	-,102		,126		-,113		,103	.05	,092
gt. häusl. Milieu	21	,158		,173		,176		,178		,148		,140	.05	,178	.05	1,77	.01	1,23	.01	1,41	.01	,179		,928
unterrichtsstör.	22	1,41		1,03		,185	.05	1,15	.01	1,20	.05	1,91		,188		1,01		,194		,197		,160		,866
unterfordert	23	,114		,115		,107		,113		,110		,106		,116		,101		,114	.05	,114	.01	,106	.05	1,10
begabt	24	,165		,175		,154	.05	,166		,131	.05	,159		,179		,164	.05	,174	.05	,169		,182		,688
selbständig	25	,164		,165		,156		,168	.01	,167		,160		,166		,153		,151	.01	,155		,167	.01	,611
sympathisch	26	1,12		1,07		,198		,196		,192		,185		1,11	.01	,198		1,11		,192	.05	,191		,993
off. Charakter	27	,146		,134		,153		,136		,147		,125		,144		,152	.05	,139		,140		,179		,395
vital	28	,150		,149	.05	,151		,161	.05	,157	.05	,157		,160	.01	,166		,146		,180		,160	.01	,579
zuverlässig	29	,182		,183		,180		,183		,183		,171	.05	,197		1,01		,185		,178		,168		,828

Tab. 4o : Mittelwerte
Signifikanztest der Skalen : WILCOXONs T; (N = 23)
Signifikanztest der Gesamtmittelwerte zwischen Meßzeitpunkten : Vorzeichentest

Man könnte also die Mittelwerte gerade zum Zeitpunkt 1 als Indikatoren eines Bezugssystems der 'Milde' und 'Strenge' der Lehrer sehen, in dem die Lehrer interessierende Skalen hervorgehoben erscheinen, andere hingegen in der jeweiligen Beurteilungssituation bedeutungsloser erscheinen. Indem wir die Gruppe der Lehrer als ganze betrachten, verfolgen wir 'Referenzfehler' (KLEITER,1977,S.39) der Lehrergruppe. Sie sollen als die Skalen definiert werden, auf denen besonders milde geurteilt wird (hohe Mittelwerte) und auf denen besonders streng geurteilt wird (niedere Mittelwerte). Man könnte vermuten, daß auf den für die Schülerbeurteilung nach Ansicht der Lehrer weniger relevanten Skalen milder geurteilt wird, da diese Eigenschaften nach Ansicht der Lehrer weniger zu Schulerfolg und -mißerfolg beitragen, als die von ihnen für relevant erachteten Eigenschaftsskalen, auf denen wir demnach eine besondere Strenge der Beurteilung erwarten können.

Zum Zeitpunkt 1 fallen besonders die Skalen 'sympathisch-', 'positive Einstellung zu Schule und Lernen-' und 'stört nicht den Unterricht-' auf; hier bekommen die Schüler gewissermaßen einen 'Vorschuß' in Form einer Zuschreibung besonders hoher positiver Ausprägungsgrade des Urteils, relativ zu Skalen wie 'kreativ-', 'akzeleriert-', 'unterfordert-', die sich durch besonders geringe Abweichungen der Mittelwerte vom O-Punkt auszeichnen.

Fassen wir nun die Mittelwerte der Skalen als Indikatoren eines Schüler allgemein betreffenden Bezugssystems auf oder liegt hier ein für den Beurteilungszeitpunkt (=Situation) spezifisches System vor ?
Im ersten Fall müßte dieses System speziell auch im Stereotyp 'des' Schülers der 1. Gymnasialklasse festgestellt werden können. Ein Vergleich der in Ränge umgewandelten Mittelwerte des genannten Stereotyps und der aktuellen Beurteilungen zum Meßzeitpunkt 1 ergab ein KENDALL -τ von 0,096 n.s.

Die Beurteilungen des Stereotyps enthalten also andere Gewichtungen der Skalen als die aktuellen Beurteilungen. Der grundsätzliche "Milde-Bonus", der übrigens bis auf wenige Skalen über alle Meßzeitpunkte anhält, ist zum Teil natürlich ein Produkt der globalen Datenverrechnung; andererseits kommen dadurch nur die wirklich prägnanten Referenzfehler der Gruppe zum Ausdruck.
Wir wollten hier aber analog zum faktorenanalytischen Vorgehen die Lehrer als Gruppe behandeln und nicht einzelfalldiagnostisch verfahren.
Die Bezugspunkte der Urteile sind also wohl eher als situationsspezifisch aufzufassen, wobei wir mit einer Veränderung rechnen müssen. Die Skalen mit den hohen Mittelwerten betreffen vorwiegend globale Aussagen über die Schüler in der Unterrichtssituation des Schuljahresanfangs; man könnte die Urteile auf diesen Skalen als einen positiven Gesamteindruck der Lehrer bezüglich Sympathie, Einstellung der Schüler zur Schule und der erlebten Disziplinsituation auffassen. Da noch kein Unterricht stattgefunden hat, sind Aussagen über so spezifische Eigenschaften wie 'kreativ-', 'akzeleriert-', 'unterfordert-' wohl kaum definitiv zu machen. Diese Skalen finden sich auch als gering relevante unter den letzten 10 % der Nennungshäufigkeiten auf Frage 1.

Eigentlich hätten wir eher erwartet, daß die Lehrer möglichst schnell eine Einschätzung der Schüler bezüglich der von ihnen als besonders relevant erachteten Eigenschaften wie 'intelligent-', 'begabt-', 'fleißig-' usw. vorgenommen hätten. Daß diese Skalen nur mittlere Abweichungen vom O-Punkt aufweisen, belegt den situationsspezifischen Charakter der Urteile; daß diese Skalen überhaupt von O abweichen, demonstriert wiederum das Wirken der IPT, in die die systematischen Interaktionen Lehrer x Urteilsvariable als Aspekt der Kategorienfehler durch ungleiche Gewichtung und Präferenzen von Urteilsvariablen eingehen, und so die kategoriale Auswahl und Gleichgewichtung verzerren (siehe: KLEITER,1977,S.41 f).

In der von uns vertretenen 4-Faktoren-Lösung war auch der wichtigste Faktor (höchster Betrag an extrahierter Varianz) der 'Disziplin'-Faktor, während der 'Intelligenz'- und 'Kontakt'-Faktor in dieser Situation von nachgeordneter Wichtigkeit waren, was nochmals die situationsspezifische Bedeutung besonders des Faktors mit dem ersten Extraktionsrang unterstreicht.

Veränderungen machen sich zwischen den Zeitpunkten 1 und 2 bemerkbar in den Beurteilungen auf den Skalen 'aufgeweckt-', 'hübsch-', 'höflich-', 'kontaktfreudig-', 'sensibel-', 'vital-' - die alle Skalen von nachgeordneter Relevanz darstellen bis auf 'aufgeweckt-', die man als einen ersten Indikator für die Beurteilung in der Dimension 'Intelligenz' werten könnte.

Zum Zeitpunkt 3 ist Interesse an einer Skala der Dimension 'Intelligenz' festzustellen ('begabt'), wo der anfängliche Bonus schwach signifikant abgebaut wird.
Weitere Veränderungen betreffen ebenfalls abbauend die Eigenschaften 'unterrichtsstörend-', während bei 'ausdauernd-', 'konzentriert-', 'kontaktfreudig-' und 'paßt sich Klasse an-' die Mittelwerte ansteigen.

Von besonderem Interesse sind die Zeitpunkte 4 und 5, zu denen die Vorentscheidung über den weiteren Verbleib der Schüler auf dem Gymnasium getroffen wurde; für besonders negativ auffällige Schüler (speziell im Leistungsbereich) wurde bereits jetzt eine Empfehlung an die Eltern ausgesprochen, ihre Kinder wieder von der Schule zu nehmen, damit die Kinder noch den Anschluß an den Lernstoff der entsprechenden Klassen der Realschule oder meistens wieder der Hauptschule noch bewältigen konnten.

Das gesamte Bezugssystem der Beurteilungssituation 4 stimmt weitgehend noch mit dem der Situation 1 überein (τ = 0,87 s.).

Die Unterschiede zum Zeitpunkt 1 hin beziehen sich vorwiegend auf Skalen des Arbeitsverhaltens ('fleißig-', 'gute Mitarbeit-', 'ausdauernd-', 'selbständig-'), deren Mittelwerte ansteigen, ebenso wie auf den Skalen 'ehrlich-', 'hübsch-', 'höflich-', 'paßt sich Klasse an-'; ob ein Schüler den Unterricht stört oder nicht, ist weiterhin wichtig, wird jedoch noch milder beurteilt als bisher. Parallel dazu entwickelt sich die Beurteilung der 'Vitalität', die, wie wir in den Faktorenanalysen oft gesehen haben, als 'störende Umtriebigkeit' betrachtet wird. Strenger wird die 'sprachliche Begabung'beurteilt.

Die Übereinstimmung des aktuellen Beurteilungssystems mit dem des Zeitpunktes 1 beträgt jedoch noch $\tau = 0.76$ s. .
Es verändert sich besonders auffällig die Skala 'Unterrichtsstörung'. Anscheinend waren die Lehrer jetzt besonders empfindlich gegen Störungen, da hier die Beurteilung im Gegensatz zu früher extrem strenger ist.Eine größere Strenge finden wir auf den Skalen 'begabt-', 'ausdauernd-', 'ehrgeizig-', 'hübsch-', 'paßt sich der Klasse an-', strenger werden die Urteile auch auf 'guter Lehrerkontakt-', 'akzeleriert-' und 'sensibel-'.

Zum Zeitpunkt 6 tritt wieder eine Beruhigung bezüglich Veränderungen ein, die allmählich zahlenmäßig aber wieder ansteigen bis zum Zeitpunkt 10, zudem sich 61 % der Beurteilungen gegenüber dem Anfangszeitpunkt verändert zeigen.
Strenger werden die Urteile in der Intelligenzdimension ('intelligent-', 'unterfordert-' ,während 'kreativ-' milder beurteilt wird), ebenfalls in der Dimension 'Sozialkontakt' ('guter Lehrerkontakt-', 'kontaktfreudig-'), die items der Arbeitsdisziplin bieten ein uneinheitliches Bild: zurückhaltender werden Fleiß, Ordentlichkeit und Ehrgeiz beurteilt, milder Mitarbeit und Ausdauer. Besonders auffällig positiv wird das 'gute häusliche Milieu-' beurteilt, das zum Zeitpunkt 6 plötzlich positiv anstieg und den signifikanten Abstand zum Eingangswert auch wahrte.

Der Zeitpunkt 11 ist wieder gekennzeichnet durch eine relative Strenge der Urteile. Es verändert sich die Beurteilung auf den Skalen im Einzugsbereich des Intelligenzfaktors ('aufgeweckt-', 'kreativ-', 'unterfordert-', 'selbständig-'), bezüglich der schulischen Disziplin bzw. Arbeitshaltung ('fleißig-', 'ehrgeizig-', 'konzentriert-', 'ausdauernd-', 'positive Einstellung zu Schule und Lernen-'), sowie auf: 'sprachlich begabt-', und den allgemeiner wertenden Skalen ohne direkte Verbindung zum Schulbereich: 'ehrlich-', 'höflich-', 'sensibel-'.

Das Präferenzsystem korreliert auch zu diesem Zeitpunkt noch signifikant mit dem eingangs des Meßzeitraumes festgestellten (τ = 0.73). Zu diesem Zeitpunkt ist aber bereits eine gewisse Angleichung an das System des Stereotyps 'des' Schülers der 1. Gymnasialklasse festzustellen (τ= 0.33 s.).

Wir haben in diesem Kapitel versucht, ein bestimmtes Milde-Strenge-System der Lehrergruppe in Anlehnung an die Faktoren der Faktorenanalyse durch die Darstellung der Mittelwerte aufzuzeigen. Die Korrelationen der von einem Faktor geladenen Variablen sind jedoch anscheinend nicht hoch genug, sodaß parallele Veränderungen der Mittelwerte dieser 'faktorentypischen' Variablen nicht zu beobachten waren. Aus der Beobachtung eines geringen Zusammenhanges zwischen der Milde-Strenge-Rangreihe der aktuell beurteilten Schüler und der Stereotypenbeurteilung 'des' Schülers (τ= 0,09) und des 'guten' Schülers (τ = 0,0025) konnte man Situationsspezifität des Systems zunächst annehmen. Es zeigte sich jedoch, daß die Rangreihe auch über eine größere Zeitspanne (Meßzeitpunkte 1 - 10, 1 - 11) in ausreichendem Maße erhalten blieb, sodaß von einer Durchgängigkeit der Beurteilungsstrenge bezüglich der Skalen gesprochen werden kann.

Unsere anfängliche Vermutung, daß die von den Lehrern als relativ irrelevant betrachteten Eigenschaftsskalen mit besonderer Milde, die relevanten mit besonderer Strenge beurteilt würden, traf nur zum Teil zu.
Sehen wir uns die Skalen mit durchgängig relativ extremen Mittelwerten an: (Tab. 41)

Skala	M (1-11)	R Streuungsbreite	Nennungshäufigkeit (Frage 1)
ehrlich	1,13	0,91 - 1,53	4
höflich	1,03	0,78 - 1,24	4
pos.Einst.zu Schule u. Lernen	0,998	0,94 - 1,30	9
sympathisch	0,99	0,85 - 1,11	-
kreativ	0,29	0,19 - 0,46	3
akzeleriert	0,16	0,09 - 0,30	2
unterfordert	0,11	0,06 - 0,19	4
sensibel	0,08	0,23 - 0,39	4

Nehmen wir als Relevanzmaßstab die Häufigkeit der Nennungen auf Frage 1: "Nach welchen Merkmalen beurteilen Sie die Persönlichkeit Ihrer Schüler?", so zeigt sich, daß die im Sinne des Bezugsschemas sehr milde und sehr streng beurteilten Skalen bis auf eine Ausnahme den von den Lehrern als irrelevant erscheinenden Skalen angehören, während die Mittelwerte der relevanten Skalen in einem Bereich durchschnittlich 'mittlerer Milde' (0,40 - 0,80) variieren.
Man könnte dieses Ergebnis so interpretieren, daß die Lehrer zwar beide Variablengruppen als relativ irrelevant betrachteten, soziale und schulmoralisch bewertende Skalen mit einem Bonus an Wertschätzung der Schülerpersönlichkeiten - unabhängig von der Leistungspersönlichkeit der Schüler - beurteilten, wobei die meisten Lehrer zu ganz spezifischen Eigenschaften wie kreativ, akzeleriert, usw. keine Beziehung fanden.

So gesehen,haben wir in den Faktorenlösungen speziell nach dem Eigenwertkriterium mit den Faktoren V - VII eigentlich 'Irrelevanz'-Faktoren extrahiert, die keine echten Dimensionen der Beurteilung darstellen, sondern eher die interindividuelle Variabilität der Urteileransichten bezüglich zu beurteilender Eigenschaften.

Die Ergebnisse dieses Kapitels erklären auch eine Erscheinung, die KLEITER (1977,S.53) darstellt, und nach der die Interkorrelationen von moralisch bewerteten Variablen unabhängig sind vom Bekanntschaftsgrad zwischen beurteilenden Lehrern und beurteilten Schülern.Die dieses Phänomen erklärende Variable ist die gleiche geringe Relevanz, die die Urteiler diesen moralisch bewertenden Variablen für den schulischen Beurteilungsprozeß zubilligen und der somit zu sehr ähnlichen Markierungen der Ausprägungsgrade durch die Lehrer-Beurteiler führt. Die von KLEITER berichteten Interkorrelationen sind dann Ausdruck einer einheitlichen Wertschätzung der Persönlichkeit der beurteilten Kinder - abgesehen von der leistungs- und schuldisziplinarischen Beurteilung, welche letztere eine Abhängigkeit vom Bekanntschaftsgrad Lehrer-Schüler aufwies (KLEITER,a.a.O., S.53).

Insgesamt sollten wir jedoch die in diesem Kapitel dargestellten Ergebnisse aus methodischen Gründen mit Vorbehalt interpretieren, nicht allein wegen der nivellierenden Wirkung der Mittelwertsberechnungen; bei der Datenverarbeitung bekommt man gerade bei großen Datenmengen die Originaldaten kaum zu Gesicht - bei einigen überprüfenden Rechnungen mit der Rechenmaschine schien sich jedoch eine Empfindlichkeit des verwendeten WILCOXON - Tests weniger gegenüber der Schiefe der Differenzen-Ver teilung, als eher bezüglich ihrer Steilheit abzuzeichnen. Gerade bei leptokurtischem Exzeß der Differenzverteilung, der häufig zu den ersten Meßzeitpunkten auftrat, erschienen manche geringen Mittelwertsdifferenzen signifikant, andere, weitaus größere Mittelwertsdifferenzen dagegen nicht.
Der fehlenden Zeit wegen wurde dieser Erscheinung nicht nachgangen.

5.9. Stabilität der Persönlichkeitsbeurteilungen über den Meßzeitraum

In der faktorenanalytischen Untersuchung wurden immer nur die Korrelationen der Skalen zu einem Zeitpunkt betrachtet. Sehen wir uns nun einmal die Zusammenhänge der Urteils-Rohwerte zwischen den Meßzeitpunkten an.

Die Korrelationen stellen hier ein Maß für die Veränderungen der Urteile auf den Skalen über die Zeit hinweg dar und geben an, wie fixiert die Urteile zu den einzelnen Zeitpunkten sind. Signifikante Korrelationsunterschiede zeigen hier, daß bedeutsame Beurteilungsveränderungen stattgefunden haben. Hypothetisch könnte man formulieren, daß mit steigendem Bekanntheitsgrad zwischen Lehrern und Schülern sich allmählich ein Schüler-"Image" bei den Lehrern verfestigt, das sich in stetig ansteigenden Korrelationen manifestiert, die schließlich einen Wert von $r = 1.00$ erreichen und diesen Wert - mit Zufallsabweichungen beibehalten. Erst dann, so könnte man definieren, "kennen" die Lehrer ihre Schüler so, daß einzelne, von diesen fixierten Einschätzungen abweichende Reize in den Beurteilungsintervallen sich kaum noch auf das Bild vom Schüler auswirken, soweit es mit unseren Rating-Skalen erfassbar ist.

Formulieren wir etwas vorsichtiger, so wollen wir also die Stabilität der Urteile über die Zeit hinweg verfolgen. Für die Beantwortung der Stabilitätsfrage bieten sich Korrelationen der Meßwertreihen pro Skala zwischen den Zeitpunkten an; wir können jedoch wieder davon ausgehen, daß die Meßwerte zu späteren Zeitpunkten (2,3,....11) von den Eingangswerten (Zeitpunkt 1) abhängig sind. Dieser Einfluß soll herauspartialisiert werden, so daß wir für Korrelation aufeinanderfolgender Meßzeitpunkte stets den partiellen Korrelationskoeffizienten $r_{n-1/n.1}$ rechnen.

In der umseitigen Tabelle der Koeffizienten fallen zunächst einmal die hohen (natürlich nicht partiellen) Korrelationen zwischen den Zeitpunkten 1 und 2 auf. Eigentlich hätten wir sehr geringe Korrelationen zwischen diesen Zeitpunkten erwartet.
Dieses Phänomen könnte , auf die festgestellte eingeschränkte Skalenbenutzung in den ersten Meßzeitpunkten und auf den in der IPT festgestellten Zusammenhang von Variablengruppen zurückzuführen sein. Bei den vermuteten Ähnlichkeitsschätzungen der vorgegebenen Eigenschaftsbegriffe genügt es dann eigentlich schon, wenn ein Urteiler nur eine Variable aus einer Gruppe interkorrelierender Variablen ebenso markiert wie zum Zeitpunkt davor; die anderen Variablen dieser Gruppe werden dann je nach der erlebten Stärke des Zusammenhanges mit dieser einen Variablen mehr oder weniger "automatisch" so markiert werden, wie zum Zeitpunkt davor.
Bei der Beurteilung von z.B. nur drei Skalenpunkten einer sieben-stufigen Skala ist die 'Treffer-Wahrscheinlichkeit' (gleich Markierung wie zum vorigen Zeitpunkt) wesentlich größer als bei der Ausnutzung der ganzen Skalenbreite. Je weiter die Skala ausgenutzt wird, desto unwahrscheinlicher sind dann aber 'Treffer', d.h. die Korrelationswahrscheinlichkeit sinkt. Bleibt sie jedoch gleich bzw. steigt sie sogar an, so ist ein systematischer Einfluß zu vermuten, der diese Tendenz der Korrelationen erklärt. Als systematischen Einfluß könnte man die steigende 'Bekanntschaft' mit dem Beurteilungsobjekt ansehen, definiert als größer werdende Stabilität der Ausprägungsgrade von Eigenschaftszuschreibungen, die vom Urteiler mit dem Namen eines Schülers in Verbindung gebracht werden.

Betrachten wir daraufhin unsere Tabelle der Korrelationskoeffizienten, so stellen wir keine durchgängige, aber eine tendenzielle Steigerung der Koeffizienten vom 2. bis zum 10. Meßzeitpunkt fest.
Die Stabilitätskoeffizienten erreichen jedoch keine solche Höhe, daß man von einem festgeschriebenen 'image' der Schüler sprechen könnte, das sich nicht mehr verändert.

Korrelationen zwischen den Meßzeitpunkten :

Variable		1 - 2	Sig	2 - 3.1	Sig	3 - 4.1	Sig	4 - 5.1	Sig	5 - 6.1	Sig	6 - 7.1	Sig	7 - 8.1	Sig	8 - 9.1	Sig	9 - 10.1	Sig	10 - 11.1	Mr
aufgeweckt	1	.57		.57	.01	.68	.05	.58	.01	.73		.70	.01	.88	.05	.84	.01	.71	.01	.47	.61
ehrlich	2	.48	.05	.58	.01	.76	.01	.58	.01	.76	.01	.45	.01	.84		.87	.05	.82	.01	.07	.55
ehrgeizig	3	.52	.05	.40	.01	.57	.05	.65		.67	.01	.79		.76	.01	.59		.55	.01	.19	.52
sprachl. begabt	4	.54	.01	.25	.01	.72	.01	.90		.88	.01	.75	.01	.99	.01	.64	.01	.79	.01	.09	.54
eig. Mitdenken	5	.56		.55	.01	.68	.01	.56	.01	.68	.01	.83	.01	.73	.05	.65	.05	.72	.01	.23	.56
fleissig	6	.68	.01	.57	.01	.69		.68	.05	.76		.79	.05	.72		.66		.67	.01	.22	.58
gt. Lehrerkontakt	7	.67	.01	.53	.05	.45		.49	.01	.72		.75	.01	.56		.63		.60	.01	.45	.53
gt. Mitarbeit	8	.71	.01	.59	.01	.63	.01	.50	.01	.74		.75	.01	.62		.67	.01	.78	.01	.35	.58
hübsch	9	.69	.01	.75		.70	.05	.62	.05	.69	.01	.48	.01	.87		.88	.01	.80	.01	.34	.62
höflich	10	.59		.55		.52		.51	.01	.68	.05	.61		.67		.71		.69	.01	.34	.53
intelligent	11	.56		.56	.01	.65		.68		.72	.01	.86		.88		.88	.01	.69	.01	.35	.62
kontaktfreudig	12	.58		.64		.51	.05	.68		.71	.05	.62	.01	.76	.01	.66	.01	.80	.01	.39	.58
konzentriert	13	.65		.66		.67		.73	.01	.85	.01	.60	.01	.76	.01	.54		.59	.01	.30	.58
kreativ	14	.57		.58		.60	.01	.41		.71	.01	.73		.78		.74		.71	.01	.26	.54
ausdauernd	15	.49	.01	.62	.05	.71		.65		.07		.69		.61		.60		.67	.01	.26	.54
ordentlich	16	.68	.05	.60	.05	.73		.64		.73		.71		.72	.01	.81	.01	.73	.01	.32	.61
paßt s. Klasse an	17	.46	.01	.71	.01	.52		.44		.46	.01	.20	.01	.46		.52	.05	.39		.35	.36
pos. Einst. Schule	18	.64	.01	.50	.01	.65	.01	.77	.01	.86	.01	.59		.64		.68	.01	.78	.01	.13	.57
akzeleriert	19	.52		.51		.56	.01	.38	.01	.52		.60	.01	.62		.79	.01	.66	.01	.40	.52
sensibel	20	.59		.57		.63	.01	.75		.71		.73	.05	.79	.01	.54		.59	.01	.43	.58
gt. häusl. Milieu	21	.86	.01	.71	.01	.80	.01	.88	.01	.97	.01	.71	.01	.61	.05	.85	.01	.75	.01	.20	.70
unterrichtsstör.	22	.66		.00	.05	.69	.01	.03	.01	.49	.01	.67	.01	.79	.05	.72		.73	.01	.40	.43
unterfordert	23	.84	.01	.55		.56		.49	.01	.63		.60	.01	.78	.01	.60	.01	.73	.01	.44	.56
begabt	24	.61	.01	.72		.67		.69		.74	.01	.63	.01	.80	.01	.51	.01	.72	.01	.09	.56
selbständig	25	.78	.01	.54	.05	.65	.05	.55	.01	.75	.01	.63	.05	.54		.51	.05	.40		.30	.52
sympathisch	26	.66	.01	.54		.60		.54	.01	.11	.01	.32	.01	.76		.74	.01	.63	.01	.31	.47
off. Charakter	27	.67	.05	.59		.59		.58	.01	.73	.05	.64		.69	.01	.50		.51	.01	.37	.53
vital	28	.69		.66	.05	.55	.01	.72		.67		.69	.01	.67	.01	.79	.01	.71	.01	.56	.63
zuverlässig	29	.60	.01	.51	.01	.65	.05	.74	.05	.68		.68		.68		.73	.01	.61	.01	.29	.56
	Mr	.625		.561		.637		.603		.651		.627		.744		.686		.680		.302	

Tab. 42 : Partielle Korrelationskoeffizienten der Zeitpunkte 2 - 11
Sig.: Signifikanz zwischen benachbarten Werten (nach FISHERs z-Transfomation der Koeffizienten)

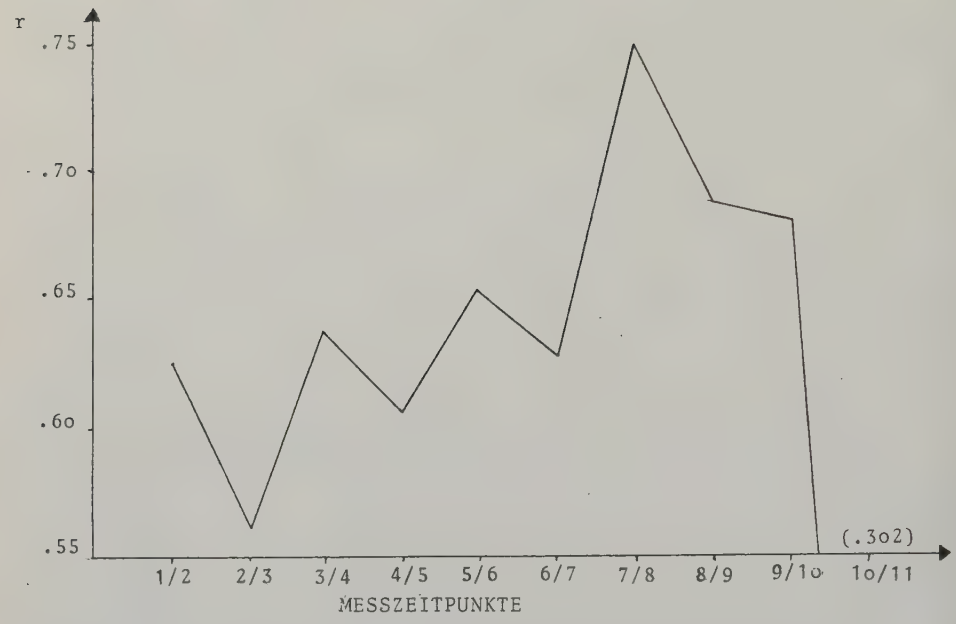


Abb.1o: Durchschnittliche Höhe der Korrelationen zwischen den Meßzeitpunkten

Die graphische Darstellung verdeutlicht das rhythmische Auftreten von steilen Anstiegen und 'Plateaus', deren leicht sinkende Tendenz mit jedem Mal flacher wird. Vor dem Hintergrund der allmählich an Höhe gewinnenden Stabilität der Urteile über die enger zusammenliegenden Meßzeitpunkte hinweg erscheint der Abfall der Stabilität vom zehnten zum - ein halbes Jahr später angesetzten - elften Meßzeitpunkt.

Da in der Betrachtung der einzelnen Skalen nach sinnvoller Gruppierung (z.B. Variablen der Leistungspersönlichkeit, Variablen schulmoralischer Bewertung, etc.) keine grundsätzlich andere interpretierbare Entwicklung der Koeffizienten festgestellt werden konnte, soll von einer Einzeldarstellung über die Entwicklung einzelner Urteilsskalen abgesehen werden.

Hier würde nur ein Verfolgen der Stabilitätsverläufe der Urteile einzelner Lehrer in einzelfalldiagnostischer Absicht oder eine Auswertung nach verschiedenen Kriterien zusammengestellter Schülergruppen Aufschluß bringen können, was allerdings den Rahmen der Arbeit erheblich ausweiten würde.

Wir wollen uns also an dieser Stelle vorerst mit der Feststellung begnügen,

- daß die Stabilität der Urteile über den Meßzeitraum eine ansteigende Tendenz aufweist;
- daß die geringe Stabilität der Urteile zwischen weiter auseinanderliegenden Erhebungen eine starke Veränderung des Beurteilungsprofils signalisiert. Leider war kein zwölfter Meßzeitpunkt mehr möglich, der einen Aufschluß über die Stabilität der Urteile am Schuljahrsende hätte geben können;
- daß die Urteile ferner nach einem halben Jahr, zu einem Zeitpunkt also, zu dem die Lehrer endgültig über den weiteren Verbleib der Schüler auf dem Gymnasium entscheiden, nicht vollständig fixiert sind. Man kann also noch nicht von einem festgeschriebenen "image" der Schüler sprechen, das sich nicht mehr verändert.

Zur Erklärung dieser Erscheinung können folgende Begründungen dienen: Einmal könnte man annehmen, daß die Lehrer ihre Schüler doch noch nicht so gut kennen, daß sie ihnen bei Nennung der Namen eindeutig Skalenwerte zuordnen können, was eine Funktion der Kontaktdauer sein könnte. Nimmt man an, daß ein Lehrer, der nur wenige Stunden in einer Klasse verbringt (etwa ein Nebenfach-Lehrer) längere Zeit benötigt, um ein "Bild" des einzelnen Schülers zu fixieren, als ein Lehrer der häufiger in der Klasse ist, so würde ein Korrelationsvergleich der Extrem-Gruppen "wenige - viele Dienststunden in der Klasse" eine Aussage über die Abhängigkeit der image-Bildung von der Kontakthäufigkeit gestatten. Eine solche Gegenüberstellung ist jedoch nicht möglich, da die Lehrer mit geringen Kontakthäufigkeiten zur Klasse zum Ziel der Homogenisierung der Stichprobe eliminiert worden waren.

Eine weitere Erklärungsmöglichkeit könnte sein, daß die Lehrer einfach über die meisten Schülereigenschaften keine Indikator-Informationen haben, die als Anhaltspunkt für eine Beurteilung dienen könnten. In diesem Fall stützt sich die Beurteilung ausschließlich auf Ähnlichkeitserlebnisse zwischen den Eigenschaftsbegriffen, zu denen Indikatoren auf proximaler Ebene beobachtet wurden (Leitvariablen), und den Eigenschaften, bei denen das nicht der Fall ist. Da diese Ähnlichkeitserlebnisse in Form der IPT über alle Meßzeitpunkte als wirksam festgestellt wurden, müßten sich diese Leitvariablen, die das Feld der Urteile vorstrukturieren, und nach denen sich die Beurteilung der anderen Variablen richtet, als 'zentrale Eigenschaften' in den Faktorenanalysen zeigen. Skalen mit eindeutig höchster Kommunalität, die sich in ihrer Höhe gut von den anderen Kommunalitäten absetzen, konnten jedoch nicht durchgängig festgestellt werden. Eine Auszählung der höchsten Kommunalitäten der 4-Faktoren-Lösung über alle Zeitpunkte ergab, daß die Skalen 'intelligent-', 'aufgeweckt-', 'begabt-', 'zuverlässig-' und 'eigenes Mitdenken' am häufigsten die höchsten

Kommunalitäten aufwiesen. Die ersten drei Skalen traten auch unter den Skalen mit der höchsten durchschnittlichen Korrelation auf. Die eindeutig höchste durchschnittliche Korrelation wies mit .75 jedoch 'gutes häusliches Milieu-' auf, über die ersten zehn Meßzeitpunkte sogar .83. Von daher könnten wir schlußfolgern, daß das 'häusliche Milieu' für die Lehrer das subjektiv am sichersten zu beurteilende Merkmal darstelle, das sich also als Kristallisationspunkt und Ausgangsbasis für die Beurteilung der anderen Merkmale anbieten würde.

Diese Interpretation gewinnt an Wahrscheinlichkeit, wenn wir bedenken, daß SCHWARZER (1976, S.75 ff) die Sozialschichtabhängigkeit der Lehrerurteile bezüglich Interesse, Sympathie, Höflichkeit, Sensibilität, Begabung in ihrer Stichprobe aufzeigte; bedenken wir auch, daß das 'häusliche Milieu' in unserer Beurteilungsreihe der Leistungsstereotypen einen eigenen Faktor und die zentrale Eigenschaft darstellte.

Dagegen spricht, daß diese Skala in den aktuellen Beurteilungen nur einmal, in der Eigenwertlösung zum Zeitpunkt 4 die 'zentrale Eigenschaft' darstellte, und daß die Korrelation der Urteile auf dieser Skala zwischen den Meßzeitpunkten 10 und 11 ebenso absinkt wie auf allen anderen Skalen, obwohl gerade diese Skala eigentlich eine der wenigen ist, die ein proximales, dem Urteiler durch Beobachtung zugängliches Merkmal darstellt.

Nun könnte man denken, daß die Urteile auf dieser Skala eine recht hohe Stabilität aufweisen, weil sich die Lehrer hier Informationen aus der Schülerkartei hätten holen können; entgegen dem sonstigen Gebrauch war aber die Schülerkartei während des Untersuchungszeitraumes den Lehrern nicht zugänglich; es könnte jedoch sein, daß die Lehrer sich diese Information durch Befragen der Schüler verschafft haben, eine solche Befragung hätte wieder die Wichtigkeit dieser Informationen für die Lehrer demonstriert.

Das Abfallen der Korrelationen vom Zeitpunkt 10 zum Zeitpunkt 11 mit der Tatsache erklären zu wollen,

- daß die Lehrer in ihren Beurteilungen 1 - 10 von falschen Annahmen über den sozio-ökonomischen Status der Schüler ausgegangen seien,
- daß gerade ab Zeitpunkt 10 eine so geringe Urteilsstabilität auftrat, nachdem die Lehrer durch die Aushändigung der Schülerkartei zum Eintragen der Zeugnisnoten offiziell über den sozio-ökonomischen Hintergrund der Schüler informiert wurden, dürfte jedoch etwas zu spekulativ sein.

5.1o. Differenziertheit der Urteile

In diesem Abschnitt sollte nicht die ganze Lehrerstichprobe untersucht werden, es sollte vielmehr eine auffällige Lehrergruppe mit einer gleich großen nichtauffälligen Gruppe verglichen werden.
Diese speziell interessierende Gruppe umfaßte die Lehrer, die sich anfänglich geweigert hatten, ihr Stereotyp 'des' Schülers zu äußern mit dem Hinweis, daß ihre Schüler zu unterschiedlich seien, um ein so globales Urteil abgeben zu können. Es waren die 6 jüngsten Lehrer bis zu 33 Jahren, die aber andererseits bei der Überprüfung von signifikanten Unterschieden der Stereotypen in Abhängigkeit vom Alter nicht aufgefallen waren, auch nicht in der Extremgruppe C 25 (genauer C 26), die genau diese Teilstichprobe umfasste.

Die Frage war, ob diese Gruppe, die eine differenzierte Wahrnehmung der Schüler verbal zum Ausdruck brachte, auch in aktuellen Urteilen eine größere Differenziertheit der Urteile aufwies. Unter Differenziertheit sollte bei der Beantwortung dieser Frage die Streuung der Urteile auf der Skala verstanden werden, nicht die Differenziertheit der Sichtweisen oder die kognitive Komplexität als Anzahl unabhängiger Urteilsfaktoren, was nachzuprüfen bei diesen kleinen Teilstichproben methodisch sowieso nicht möglich gewesen wäre.

Diese 6 Lehrer hatten auf Frage 1 durchschnittlich 7 mögliche zu beurteilende Merkmale von Schülern genannt gegenüber durchschnittlich 5 Nennungen der Stichprobe der 6 ältesten Lehrer, woher die Vermutung einer ebenfalls differenzierteren Beurteilung auf den vorgegebenen Skalen resultierte.
Als Vergleichszeitpunkt wurde der Meßzeitpunkt 11 gewählt, zu dem man die im Meßzeitraum beste Kenntnis der Schüler vermuten kann. Als Prüfverfahren wurde der F-Test gewählt. Bei df1= 5 und df2= 5 ergaben sich keine signifikanten Differenzen zwischen den Lehrern, weswegen von einer Darstellung der nun überflüssigen Ergebnisse abgesehen werden soll.

Man könnte dieses Ergebnis dahingehend interpretieren, daß diese Untergruppe der Lehrer keine differenziertere Beurteilung der Schüler vornahm, wobei allerdings in Rechnung zu stellen ist, daß die Differenziertheit hier definiert wurde als Streuung der Urteile auf vorgegebenen Skalen.

VI Diskussion der Ergebnisse

Die grundsätzliche Fragestellung war, ob in den Urteilen der untersuchten Lehrer über ihre Schüler Zusammenhänge der verwendeten Eigenschaftsbezeichnungen festzustellen waren und auf welche Faktoren diese eventuellen Zusammenhänge zurückzuführen seien.
Aus den Faktorenanalysen der Urteile über die Stereotype "der" Schüler, "der gute Schüler" und "der schlechte Schüler" resultierten fünf Faktoren, die zum Teil schon aus der Literatur bekannt sind:

I : Geistige Regsamkeit
II : Schuldisziplin
III : Kontaktoffenheit
IV : psycho-physische Konstitution
V : häusliches Milieu

Nur die drei ersten Faktoren waren signifikant, weswegen also die aus den Schülerstereotypen gewonnene Urteilsstruktur (IPT) als dreidimensional angesehen wird.

Der Faktor IV, der etwa dem Faktor "Empfindsamkeit vs. Robustheit" von HOFER am nächsten kam, klärte nur noch einen geringen Teil der Varianz auf. Schon bei HOFER ist dieser Faktor als recht schwach zu beurteilen, obwohl er größere Varianzanteile bindet. Bei der Anwendung der Prüfgröße a^2/h^2, die FÜRNTRATT (1969) anregt, stellt er sich nämlich als ein 'doublet' (Faktor mit nur zwei wesentlichen Ladungen; PAWLIK,1968,S.266) dar, in dem nur bei der Skala 'kompliziert - unkompliziert' ein nennenswerter Teil ihrer Kommunalität durch den Faktor aufgeklärt wird, während die Skala 'ausgeglichen - launisch' das Kriterium der Prüfgröße (.50) nicht erreicht, die Skala 'sensibel - robust' nur knapp. Autoren wie MASENDORF und SCHWARZER (1976) haben in der angestrebten Replikation der 5-Faktoren-Lösung ebenfalls Schwierigkeiten, diesen Faktor zu halten.

Der Faktor V bindet noch weniger Varianz als Faktor IV, zudem ist es ein 'single' und scheint überdies noch ein Methodenfaktor zu sein, da er in keiner der immerhin 34 durchgeführten Faktorenanalysen jemals wieder auftauchte. Bei der Überlegung, welche methodische Besonderheit ihn hervorgerufen haben könnte, stoßen wir auf die Fragen 4, 5 und 6 (s.S.74 ff) die das Material für die vorliegende Analyse lieferten. In zwei der drei vorgegebenen Stereotypen, die beurteilt werden sollten, hatten wir gewissermaßen eine Normierung der (fiktiven) Reizobjekte durchgeführt, denn 'der gute Schüler' und 'der schlechte Schüler' stellen ja Extreme eines Merkmals dar, nämlich der Schulleistung. Die Variable 'häusliches Milieu' weist mit Abstand die größte Kommunalität auf (was allerdings bei 'singles' häufig ist) und wird vom Faktor 'geistige Regsamkeit' zu .37 geladen, von 'Schuldisziplin' zu .48, sie korreliert zudem mit allen anderen Variablen positiv, am höchsten mit 'intelligent', 'fleißig', 'ordentlich', 'positive Einstellung zu Schule und Lernen', 'höflich'. Man könnte also sagen, daß die Lehrer von dieser Variablen bis zu einem gewissen Grad zumindest auf Variablen der Faktoren 'geistige Regsamkeit' und 'Schuldisziplin' schließen. Allerdings ist wie gesagt die Güte dieses Faktors zu gering, um diese Schlußfolgerung mit Nachdruck vertreten zu können (s.dazu auch HÖHN, 1967, S.73).

Die nächste Fragestellung betraf die Faktoren aus aktuellen Urteilen. Dazu ist es notwendig, die Informationssituation der Urteiler zu berücksichtigen, die sich ja in dieser Studie über die Meßzeitpunkte hinweg veränderte.
Nach dem Modell von BRUNSWIK (1956) über die Attribuierung distaler Merkmale (Eigenschaften,latente Verhaltensdispositionen) , bedient sich der Urteiler verschiedener Indikatoren (proximaler Merkmale des Beurteilten), von denen er subjektiv eine Merkmals-Validität annimmt, da er die Eigenschaften selbst ja nicht wahrnehmen kann. Diese Indikatoren sind nach BRUNSWIK direkt beobachtbare Merkmale des Beurteilten wie Verhaltensweisen, Verhaltensprodukte, physisch-gegenständliche Merkmale (Geschlechtszugehörigkeit, Aussehen, etc.) und sozio-ökonomische Gegebenheiten.

Es werden also Informationen bezüglich dieser proximalen Merkmale der Schüler vorsätzlich (mit dem Ziel, Eigenschaften der Schüler zu rekonstruieren) oder unabsichtlich gesammelt. Ob solche Informationen vorliegen oder nicht, können wir erst aus der Reaktion des Urteilers - für uns das Ankreuzen eines bestimmten Ausprägungsgrades auf einer der vorgegebenen Skalen - rückschließen.

Sehen wir uns daraufhin einmal die Informationssituation des Lehrers zum Zeitpunkt 1 an. Hier kann er allein schon von der Zeit her, die er in der Klasse verbracht hat, nur wenige Informationen gesammelt haben. Und wenn er aktuelle Informationen hat, so ist noch fraglich, ob er sie dem richtigen Schüler zuordnen kann, der ja auf der Liste nur durch seinen Namen repräsentiert ist. Zudem hat der Untersucher eine sonst nicht übliche Deprivation der Lehrer bezüglich der Vorinformationen betrieben, und da es sich um Schüler der 1. Klasse handelte, liegen auch keine Informationen aus irgendeiner anderen Quelle vor. Da er aber andererseits in der Zwangslage ist, ein Urteil abgeben zu müssen, könnte man der Ansicht sein, der Lehrer würde sein Urteil aufgrund bloßen Ratens abgeben und sich dieser Urteile demgemäß nicht sicher sein.
Die Faktorenanalyse der Urteile des Zeitpunktes 1 widerspricht jedoch dieser Annahme, da sich von der Analyse der Stereotypen-Beurteilungen her bekannte Faktoren ergaben:

- fünf Faktoren (nach Eigenwert-Kriterium)
- die ersten drei Faktoren waren signifikant:
  geistige Regsamkeit, Schuldisziplin, Kontaktoffenheit
- als Abweichung: der nicht signifikante Faktor 'häusliches Milieu' aus der Stereotypenbeurteilung wurde nicht reproduziert und durch den Faktor 'spezielle Begabung' ersetzt, der ebenfalls ein 'single' ist. Auch dieser Faktor könnte zusätzlich ein Methodenfaktor sein; er lädt die Markiervariable 'eher sprachlich begabt - eher mathematisch begabt'. Denken wir daran, daß 61 % der Stichprobe Fachlehrer für Deutsch, Fremdsprachen und Mathematik waren, die bis auf einen in die Fragen 1 und 3 ihre fachspezifischen Vorstellungen einbrachten und wohl ebenso in die Beurteilungen, so könnten wir beinahe von einem 'Sprachlehrer - Mathematiklehrer-Faktor' sprechen, der bei einem bedeutsameren Varianzbetrag ein Indiz für eine fachspezifische Sichtweise von Untergruppen der Stichprobe hätte sein können.
- die Kongruenzkoeffizienten lagen mit .99, .99, .96, .93, .66 sehr hoch, sodaß man wenigstens bei den signifikanten Faktoren von fast vollständiger Identität sprechen kann.

Dieses Ergebnis kann sicher nicht auf bloßes Raten beim Ausfüllen der Beurteilungsbogen zurückzuführen sein. Die Lehrer müssen beim Urteilen höchst ähnliche Strukturisierungskriterien verwendet haben wie bei der Beurteilung der Stereotypen. Die Anwendung dieser Kriterien geht anscheinend mit einem Gewinn an subjektiver Urteilssicherheit einher, da die von den Lehrern angegebene Urteilssicherheit erheblich von totaler Unsicherheit abweicht. Letzte Klärung dieser Vermutung könnten jedoch nur Kontrolluntersuchungen bringen. Es könnte nämlich sein, daß bereits kleinste Informationen von den Reizpersonen vorliegen (Kleidung, Aussehen, 'Klassenklima', etc.), die ausreichen, um wenigstens einige distale Merkmale (Eigenschaften) zuzuschreiben, die sich wiederum als 'primacy'-Effekt erzeugende Variablen etabliert haben.

Zur Identifizierung dieser schon zum Zeitpunkt 1 mit Information besetzten Variablen könnten wir uns unsere Beobachtung zunutzemachen, daß über den Meßzeitraum mit höherem Informationsstand die subjektive Urteilssicherheit und das Abweichen der markierten Ausprägungsgrade der Eigenschaftsskalen kontinuierlich parallel ansteigen und zu .90 korrelieren. Angewandt auf den Zeitpunkt 1 müssten dann die Variablen mit den größten durchschnittlichen Abweichungen vom Mittelpunkt der Skala die 'informationsbesetzten' Variablen darstellen. Bei den in unserem Fall sehr steilen Häufigkeitsverteilungen der Markierungen auf den Skalen können wir davon ausgehen, daß die Skalen mit den höchsten Mittelwerten identisch sind mit den Skalen, die die größten durchschnittlichen Abweichungen der Markierungen von den Skalenmittelpunkten darstellen. Diese Variablen waren 'ehrlich-', 'höflich-', 'positive Einstellung zu Schule und Lernen-' und 'sympathisch-', von denen festgestellt worden war, daß ihnen die Lehrer eine relativ geringe Relevanz für die Beurteilung von Schülern zugebilligt hatten. Man könnte mit einiger Vorsicht daraus schließen, daß sich das erste Urteil, das zudem durchgängig positiv war, vorwiegend auf sozial und schulmoralisch bewertende Variablen (KLEITER,1977,S.52) geringer Relevanz bezieht. Relativiert wird diese Aussage

dadurch, daß diese Variablen nicht auch die größten Kommunalitäten in der Faktorenanalyse des Zeitpunktes 1 aufwiesen, das heißt in keinem umfangreicheren Ausmaß andere Eigenschaften implizierten.

Es muß hier jedoch nochmals darauf hingewiesen werden, daß die unstrukturierte Situation, in der die Anwendung der IPT festgestellt wurde, nicht im Klassenzimmer stattfand, sondern vor der Mappe mit den Beurteilungslisten, und daß wir die Replikation der IPT provozierten, indem wir auf die Lehrer Druck ausübten, Schüler zu beurteilen, über die sie gar keine oder fast keine Informationen hatten; zu diesem Zeitpunkt 1 wird auch eine gewisse Unnatürlichkeit dieser Untersuchungssituation klar, wie sie ebenfalls bei PASSINI u. NORMAN (1966) zu finden ist. Um den Informationsstand der Lehrer zu diesem Zeitpunkt überprüfen zu können, hätten wir - aus BRUNSWIKS's Modell abgeleitet - maximal vereinzelte Informationen auf dem Niveau proximaler Merkmale annehmen dürfen und ihnen keine Beurteilungslisten mit Eigenschaftsbegriffen, sondern mit items in der Form von Verhaltensweisen, physischen Merkmalen, etc. vorlegen müssen. Wir kommen hier von einer anderen Seite her zum Erkennen der Notwendigkeit der Verwendung von Verhaltensitems, ähnlich, wie sie SCHWARZER/SCHWARZER (1977) vorschlagen.

Nun dienen solche Orientierungshilfen wie die IPT zwar dazu, unstrukturierte Situationen zu strukturieren, sich schnell zurechtzufinden und Verhaltensunsicherheiten abzubauen, mit fortschreitender Kontakthäufigkeit jedoch müsste die Beurteilungssituation für die Lehrer strukturierter werden mittels der größer werdenden Zahl der Informationen über die Schüler einerseits und dank der Strukturierungshilfe der IPT andererseits.
Welche Entwicklung nimmt nun die IPT mit zunehmender Orientiertheit und (Urteils-)Verhaltenssicherheit der Urteiler?

Berechtigen also die rechnerischen Befunde der einzelnen Meßzeitpunkte eher zu einer idealistischen oder realistischen Interpretation der IPT?

Eine idealistische Interpretation der IPT wäre indiziert, wenn eine Ähnlichkeit der kognizierten und perzipierten Zusammenhänge von Eigenschaften in der Form festgestellt wird, daß sich die Faktoren aus der Beurteilung fiktiver Schüler (Stereotypen) auch aus den Beurteilungen realer Schüler replizieren lassen. Eine realistische Interpretation wäre angezeigt, wenn die Faktoren sich inhaltlich veränderten oder überhaupt nicht mehr aufzufinden wären, sodaß eine isomorphe Widerspiegelung "realer" Kovariationen von Verhaltensdispositionen (s.dazu ULICH/MERTENS,1973,S.115) angenommen werden kann. Dieses Modell geht allerdings von der Annahme aus, daß diese Verhaltensdispositionen (Eigenschaften) "in Wirklichkeit" "in" den Schülern nicht zusammen auftreten.
Die Vorstellung einer möglichen realistischen Interpretation der IPT ist jedoch schon von der Wahrnehmungstheorie her indiskutabel, da eine isomorphe Widerspiegelung realer Gegebenheiten speziell in Untersuchungen mit hoch inferenten items nicht erwartet werden kann. Durch Verwendung von items auf 'Eigenschafts'-Niveau haben wir Beurteilungstendenzen gewissermaßen schon vorprogrammiert - und haben sie auch zu jedem Meßzeitpunkt in Form relativ hochkongruenter Faktoren wiedergefunden.

Zu einzelnen Zeitpunkten (4,5,11) haben wir jedoch eine inhaltliche Kontamination der Faktoren I ("Geistige Regsamkeit") und II ("Schuldisziplin") in den Faktorenmatrizen feststellen müssen.

Hier wurden gravierende Unterschiede der Ergebnisse von Faktoren-Identifikation und -Interpretation gefunden, die dazu führten, daß die zu diesen Zeitpunkten entstandenen Faktoren mit den kontaminierten Inhalten der Faktoren I und II der IPT anders benannt werden müssten ("Schuleignung"). Zu diesen kritischen Zeitpunkten war das Wiederauffinden der IPT-Faktoren als Faktoren auch der aktuellen Urteile nicht eindeutig zu belegen. Eine solche Kontamination der Faktoren "Begabung" und "Arbeitshaltung" beobachtete auch SCHWARZER (1977,S.61), die

Schüler der ersten Gymnasialklasse nach einem halben Schuljahr von ihren Lehrerinnen beurteilen ließ.
Während bei SCHWARZER diese Kontamination nur an die 3-Faktoren-Lösung geknüpft war, trat diese Erscheinung in der vorliegenden Arbeit bei der Eigenwert-, der 4-Faktoren-Lösung sowie bei schiefwinkliger Rotation stets zu denselben Zeitpunkten 4, 5 und 11 auf, weswegen die Annahme einer Lösungsspezifität diese Erscheinung nicht erklären kann. Bedenkt man dagegen neben der Kontamination der Faktoren I und II die geringere Korrelationshöhe der Skalen zu diesen Zeitpunkten und die gegenüber der ursprünglichen IPT größere Faktorenzahl, so liegt als Interpretation der Summe dieser Erscheinungen eine differenziertere Beurteilung bei gleichzeitig weniger intensiver'Anwendung'der IPT und der Beurteilung nach einem 'Geistige Regsamkeit' und 'Schuldisziplin' übergreifenden '-Gymnasialeignungs-'Kriterium näher. Unterstützt wird diese Interpretation von der zeitlichen Koinzidenz des Phänomens mit weitreichenden Entscheidungen der Lehrer über die weitere Schullaufbahn der Schüler.
Ursprünglich war hier die Meinung vertreten worden, daß der von einem Faktor gebundene Varianzbetrag eine Informationsquelle über die Wichtigkeit dieses Faktors für den Beurteiler darstellen könne; in der Gegenüberstellung der Ergebnisse der Faktorenextraktion nach Eigenwertkriterium und der 4-Faktoren-Lösung würde dieser Auswertungsgesichtspunkt jedoch in Frage gestellt: der Extraktionsrang bzw. der Betrag an gebundener Varianz eines Faktors stellte sich als abhängig von der Entscheidung des Autors bezüglich der Zahl der zu extrahierenden Faktoren heraus.
Da die Zahl der resultierenden Faktoren letztlich eine willkürliche Entscheidung eines Untersuchers ist, die zudem von der Eindeutigkeit der verwendeten Abbruchkriterien abhängt, halten wir es nicht für statthaft, wie HOFER (1969,S.86) zu behaupten: "Die implizite Theorie von Lehrern in bezug auf die Persönlichkeit von Schülern zeigt eine fünf-dimensionale Struktur"; ebenso kann also die Rolle, die einzelne Faktoren in der Schülerbeurteilung spielen, nicht eindeutig erfasst werden, da sich die 'Wichtigkeit' der Faktoren, definiert als

der Betrag der durch den Faktor gebundenen Varianz, ebenfalls die von der Zahl der extrahierten Faktoren abhängig zeigte.

Um zu einer Abschätzung der Gültigkeit solcher und ähnlicher Untersuchungen für den gemeinten Sachverhalt der Urteilstendenzen in Lehrerurteilen zu kommen, sollen abschließend noch einmal einige besonders kritische Punkte in Anlage und Auswertung der Studie angesprochen werden, wobei wir uns zum Teil auf die Darstellung der Kritikpunkte von GIGERENZER u. STRUBE (1978) beziehen.

- Zunächst wird angenommen, daß Adjektive hinreichend geeignet sind, um die implizite Persönlichkeitstheorie von Lehrern und die Lehrerurteile über Schüler zu repräsentieren. HUBER u. MANDL (1979) fanden dagegen, daß in 'naturalistischen' Persönlichkeitsbeschreibungen die Lehrer ihrer Stichprobe sowohl Eigenschaftsbezeichnungen als auch Wortsequenzen über psychische Zustände und Verhaltensweisen ("durch Erfolge motiviert", "rennt immer herum") und Sätze ("sie hat die negative Einstellung ihrer Mutter der Schule gegenüber übernommen") verwendeten (a.a.O.,S.6.). In Voruntersuchungen hatte der Autor ebenfalls unstrukturierte Beurteilungen erhoben. Es zeigte sich jedoch, daß ein großer Anteil der zu einem Zeitpunkt verwendeten Urteilsvariablen zu einem zweiten Beurteilungszeitpunkt nicht mehr verwendet wurde und daß dafür neue Variablen durch die Lehrer eingeführt wurden, was methodisch nicht zu bewältigen war. Außerdem blieb zwar die Datenbasis frei von methodischen Vorgaben des Untersuchers, diese untersucher-spezifischen Einflüsse traten dann jedoch in der notwendigen inhaltsanalytischen Bearbeitung der stetig wechselnden items in Erscheinung.

- Mit GIGERENZER u. STRUBE (1978,S.78) sind wir der Meinung, daß vom Untersucher vorgegebene items bereits die in der Untersuchung 'gefundene'Urteilsstruktur vorbestimmen. Die logische Folge davon ist, daß die errechnete Faktorenstruktur eher die Struktur des Meßinstrumentes darstellt als die kognitiven Strukturen von Lehrern. Um wenigstens eine artifizielle Auswahl der verwendeten Eigenschaftsbegriffe zu vermeiden, waren in dieser Arbeit für die Lehrer beurteilungsrelevante Eigenschaften von diesen selbst erfragt worden. Zu einer Reduktion dieser Information waren wir jedoch schon aus Gründen des geringeren Zeitaufwandes für die Vpn gezwungen. Eine Anzahl von Skalen, die nur wenigen Lehrern als relevant erschienen, mußte ausgeschieden werden, obwohl dadurch die Chance vergeben wurde, Variabilität der IPTn abzubilden; da jedoch schon in der Literatur (HOFER, 1969; COHEN,1969; HANKE,LOHMÖLLER,MANDL,1978) der Nachweis urteilsspezifischer IPTn nicht gelungen war, hatten wir in der Auswertung auch keine Q-Analyse geplant. Im Nachhinein erschien es nach den Ergebnissen einer Arbeit von HUBER u.MANDL (1979) schon eher wahrscheinlich, daß urteilerspezifische IPTn bzw. Urteilsstrukturen anhand des vorliegenden Datenmaterials hätten nachgewiesen werden können. Die genannten Autoren hatten ebenfalls nur von den Lehrern verwendete Beurteilungskategorien in die Untersuchung einbezogen, die jedoch durch die Erhebung unstrukturierter,'naturalistischer' Schülerbeschreibungen eine noch größere 'Lebensnähe' aufwiesen als die in der vorliegenden Untersuchung verwendeten Kategorien.

  Da für den Untersucher jedoch keine Zugangsmöglichkeit zum Rechenzentrum mehr bestand, mußte ein Q-sort der Beurteilungsbogen unterbleiben.
  Die mit unserer Erhebungsmethodik gewonnene Liste beurteilungsrelevanter Eigenschaftsbegriffe unterscheidet sich im Wortlaut in weiteren Teilen kaum von den in anderen Untersuchungen gefundenen oder erstellten Eigenschaftslisten (HÖHN,1967; HOFER,1969), was für eine rollenspezifisch durchgängige Verwendung dieser Begriffe spricht.

- Neben der Art der verwendeten Eigenschaftenitems wirkt auch die Anzahl der items als Determinante der resultierenden Faktorenstruktur (IPT). HOFER (1969,S.86) stellte eine Korrelation von r = .90 zwischen der Zahl der verwendeten Eigenschaften und der Zahl der extrahierten Faktoren in neun Untersuchungen zur IPT fest. In der vorliegenden Arbeit wurde zwar stets die gleiche Anzahl von Eigenschaftsbegriffen verwendet, zu verschiedenen Zeitpunkten wurde nach der Eigenwertlösung jedoch eine unterschiedliche Anzahl von Faktoren (5-7) extrahiert. Wir würden diese unterschiedlichen Faktorenzahlen gern als situationsspezifisch interpretieren; da die meisten dieser Faktoren aber methodisch auf schwachen Beinen stehen, soll von einer Interpretation Abstand genommen werden. Obwohl manche dieser Faktoren durchaus sinnvoll erscheinen, könnte es dennoch sein, daß sie ihre Existenz nur der Unschärfe der Abbruchkriterien (flacher Verlauf der Eigenwertkurve bei $\lambda = 1$, keine deutliche Diskontinuität an der Stelle $\lambda = 1$) verdanken, sodaß wir leicht der Gefahr einer Überinterpretation unterliegen.

- Zur Vervollständigung der Skalen wurde bisher (z.B. HOFER,1969), das Gegenteil der ausgewählten Adjektive vom Untersucher selbst gebildet, wobei implizit angenommen wird, daß die vom Untersucher hinzu konstruierten Gegensätze identisch sind mit den von den Lehrern kognizierten und daß durch diese Gegenteile die Bedeutung der Adjektive nicht verändert wird. Damit wird die lexikalische mit der psychologischen Wortbedeutung gleichgesetzt. In der vorliegenden Untersuchung wurde diese Skalierung von den Lehrern selbst vorgenommen, was manchmal zu überraschenden Ereignissen führte.
Erstaunlich war die relativ große Deckungsgleichheit der zu den Adjektiven genannten sprachlichen Gegenteile , die nach der Ansicht von KELLY (1955) auf eine weitgehende Ähnlichkeit der konnotativen Bedeutung der verwendeten Begriffe hindeutet. Es scheint sich also bei diesen Begriffen um Beurteilungskategorien zu handeln, die in der untersuchten Lehrergruppe weitgehend gleich verstanden werden.

Ob diese Bedeutungsähnlichkeit eine rollenspezifische ist, kann wegen der fehlenden Vergleichsmöglichkeit mit anderen Stichproben nicht entschieden werden.

- Ein weiterer Aspekt der Validitätsbeurteilung ist die Stichprobenauswahl. In der Vorbildarbeit liegt der Erhebungsmethodik der aktuellen Beurteilungen nach eine nach dem Kriterium der Teilnahme-Motivation ausgelesene Stichprobe vor. In unserer Studie konnten alle Lehrer der untersuchten Schulen zu einer Teilnahme bewogen werden. Nur diejenigen Lehrer waren aus der Stichprobe genommen worden, die nur wenige Stunden pro Woche in einem Nebenfach in den Klassen verbrachten, um nicht noch eine weitere Variable in die Untersuchung einzubringen. Die fast totale Ausschöpfung der Population der Gymnasiallehrer am Untersuchungsort und die Homogenisierung der Stichprobe kann jedoch nicht darüber hinwegtäuschen, daß gerade bezüglich der Teilnahme-Motivation eine gewisse Heterogenität in der Stichprobe herrschte, die sich in einer unterschiedlichen Durchführungsgüte der Beurteilungen niederschlagen könnte.

- Die nach unserem und dem Verfahren der Vorbildarbeit errechnete Faktorenstruktur stellt notwendigerweise eine "mittlere" Urteilsstruktur dar. Ist die Stichprobe bezüglich einer oder mehrerer verdeckter, aber eine ganz bestimmte Urteilsstruktur nach sich ziehender Urteiler-Variablen inhomogen, so entspricht die von uns errechnete Faktorenstruktur möglicherweise keiner der Faktorenstrukturen der Stichproben-Untergruppen.
  Da jedoch weder in der Vorbildarbeit noch in anderen Untersuchungen (z.B. HANKE,LOHMÖLLER u. MANDL,1978) spezifische IPTn von Stichprobenuntergruppen festgestellt worden waren, bestand keine Veranlassung, an der IPT als rollenspezifischem Stereotyp zu zweifeln. Wie bereits gesagt, wurde aber von HUBER u. MANDL (1979) die Existenz von Urteilertypen aufgezeigt.

- Die nächste verwendete Implikation ist ebenso gebräuchlich wie problematisch, nämlich die Annahme, das Lehrerurteil habe das Informationsniveau einer Intervallskala. Gleich großen Skalenpunktunterschieden werden hierbei gleich große Unterschiede der psychologischen Bedeutung der Urteile auf diesen Skalen unterstellt. Aus der Untersuchung der Skalenausnutzung in Abhängigkeit von der subjektiven Urteilssicherheit könnten wir schlußfolgern, daß bei größerer Unsicherheit der Urteiler die Skalenpunkte 0 und 1 in ihrer psychologischen Bedeutung eine größere Distanz aufweisen als im Zustand größerer subjektiver Sicherheit. Neben den skaleninternen Bedeutungsunterschieden der Skalenpunktdistanzen könnten auch die unterschiedlichen Mittelwerte der einzelnen Skalen gerade zum Zeitpunkt 1 als Bedeutungsunterschiede der Skalenpunktedistanzen in Abhängigkeit von der Skalenbenennung interpretiert werden. Demnach würden die Vpn auf von ihnen als weniger beurteilungsrelevant eingeschätzten Skalen mit schulmoralisch bewertenden Benennungen die Skalenpunkte als enger zusammenliegend erleben, sodaß sie, um ihr Urteil zum Ausdruck zu bringen, weiter vom Skalenmittelpunkt entfernte Skalenpunkte ankreuzen; bei der Beurteilung auf Skalen der Leistungspersönlichkeit dagegen scheinen die Skalenpunktdistanzen größere Bedeutungsunterschiede zu beinhalten.

- Eine Implikation des verwendeten mathematischen Zusammenhangsmaßes der Korrelation stellt die Annahme nur linearer Zusammenhänge der Beurteilungskategorien dar.
Ein konstanter Zusammenhang hoher Ausprägungsgrade einer Skala mit mittleren Ausprägungsgraden einer anderen Skala wird nicht erfaßt (z.B. "weder sehr intelligentenoch sehr unintelligente Schüler sind sehr sympathisch; sehr sympathisch sind vielmehr durchschnittlich intelligente Schüler"). Urteilszusammenhänge dieser Art werden durch das verwendete Maß des Korrelationskoeffizienten von vornherein ausgeschlossen. Daraus resultieren inhomogene Verteilungen der Urteile auf den Skalen, deren Folgen für die Faktorenanalyse im nächsten Abschnitt angesprochen werden sollen.

- Vorausgesetzt wird nämlich bei der gleichen rechnerischen Behandlung der Skalen eine homogene Verteilung der Urteile auf den Skalen. Bei den Mittelwertsberechnungen mußten wir jedoch feststellen, daß die Verteilungsformen unterschiedlich waren. Zwar wurden zur Vermeidung einer Umfangsausweitung der Arbeit nur stichprobenartige Rechnungen angestellt, es deutete sich jedoch eine Abhängigkeit der Verteilungsform der Urteile auf den Skalen von der subjektiven Urteilssicherheit der Vpn und der Einschätzung der Beurteilungsrelevanz einzelner Skalen durch die Vpn an. Das äußere Anzeichen dafür war eine geringere oder weitere Ausnutzung der Skalen. Eine geringere Streuungsbreite der Urteile wies demnach auf eine steilere Verteilung hin, eine größere Streuungsbreite auf eine weniger steile Verteilung der Urteile auf der Skala.
  Diese inhomogenen Verteilungen führen zu einer Unterschätzung des tatsächlichen Zusammenhangs der Urteile und dadurch zu einer artifiziellen Erhöhung der Faktorenanzahl. Die Vermutung, daß nach der Eigenwertlösung zu viele Faktoren extrahiert werden (GUTTMAN's Eigenwertkriterium gibt ja nur die obere Grenze der extrahierbaren Faktorenzahl an), war bereits im Kapitel 'Die Eigenwertlösung' an Ort und Stelle angesprochen worden und hatte zur Durchführung der 4-Faktoren-Lösung geführt, die nach ihren Parametern (Anzahl der Variablen pro Faktor, deren Kommunalität durch den Faktor bedeutsam aufgeklärt wird; Prozentsatz der aufgeklärten Kommunalität) die best-interpretierbare Lösung zu sein schien. Vor der Durchführung einer Faktorenanalyse sollte also die Homogenität der Verteilungen der Urteile auf den Skalen überprüft werden.

- Als weitere Informationsquelle werden die unterschiedlichen Mittelwerte der Skalen nur in wenigen Untersuchungen (SCHWARZER,1976; COHEN,1969) herangezogen.

  Der Grund dafür liegt in der moderneren Definition der IPT als Faktorenstruktur der Urteile; durch die Faktorenanalyse bekommt man den ganzen Datensatz auf einmal in den Griff.

Andererseits wurde es durch den Verzicht auf die Errechnung anderer Kennwerte aus dem Datenmaterial unmöglich , fundierte Hypothesen über beurteiler- oder beurteiltenspezifische IPTn zu erstellen, da die durchgeführten Analysen nach dem Q-Modus keine unterschiedlichen Urteilertypen feststellten, andere Analysen, z.B. nach dem T-Modus unseres Wissens nicht durchgeführt wurden. Mittelwerts- sowie Varianzunterschiede aber werden in Faktorenanalysen als Fehler betrachtet und wegstandardisiert.

Die Stichprobe der vorliegenden Arbeit war jedoch zu klein, um mit den herkömmlichen statistischen Verfahren Mittelwertsvergleiche anstellen zu können. Die angestellten globalen Mittelwertsvergleiche der einzelnen Skalen zwischen den Meßzeitpunkten zeigten ein uneinheitliches Bild, weswegen von - wahrscheinlich recht artifiziell ausfallenden - Interpretationsversuchen Abstand genommen wurde. An Gruppenvergleichen extremer Mittelwerte konnte festgestellt werden, daß von den Lehrern für die Schülerbeurteilung als wenig relevant betrachtete Skalen mit schulmoralisch bewertenden Eigenschaftsbezeichnungen über alle Meßzeitpunkte hinweg besonders milde beurteilt wurden.

Der letzte Diskussionsbeitrag bezieht sich auf die Interpretation der Ergebnisse von Faktorenanalysen. Durch die von HAYS (1958) angeregte Anwendung der Faktorenanalyse auf den Untersuchungsgegenstand des Urteilsverhaltens wird impliziert, daß Faktoren psychologisch sinnvolle Konstrukte über das Urteilsverhalten darstellen. GIGERENZER u. STRUBE (1978,S.81) sprechen in diesem Zusammenhang von der Faktorenextraktion nach dem Eigenwertkriterium und der anschließenden Varimax-Rotation als von einer "Ritualhandlung" der forschenden Psychologen. Allgemein richten sich die Interpretationsansprüche der Ergebnisse nach der Präzision der mathematischen Berechnungen, nicht, wie eigentlich notwendig, nach der Zuverlässigkeit der Eingangsdaten.

Durch diese Gewohnheit bekommen die in einer Faktorenanalyse errechneten Faktoren den Charakter unumstößlicher Grundtatsachen 'im Menschen'. Dabei wird bereits durch Vorgabe von Fragebogen, Schätzskalen oder semantischen Differentialen die Informationsdifferenzierung vorab festgelegt ohne Berücksichtigung der individuellen Beurteilungspraxis der Vpn. Auf dieses methodisch determinierende Differenzierungsmodell folgt mit der Anwendung der Korrelationsstatistik und Faktorenanalyse ein mathematisch determinierendes Interpretationsmodell der Daten, das "bereits durch seine Anwendung auf einen psychologischen Gegenstandsbereich G eine psychologische Theorie über diesen Gegenstandsbereich G impliziert." (GIGERENZER, 1978,S.110).

Zur Thematik der Auswirkungen unterschiedlicher Erhebungsmethoden auf die resultierenden Faktoren liegt eine Untersuchung von HUBER u. MANDL (1979) vor.

> Die Autoren lassen eine Stichprobe von 26 Junglehrern 7 Schüler aus in drei Leistungsniveaus eingeteilten Klassen in unstrukturierten Interviews beschreiben. Die Interviews werden inhaltsanalytisch ausgewertet, indem zur Datenreduktion bedeutungsgleiche Beschreibungen zusammengefasst und von den Untersuchern kategorisiert wurden.
> Die Faktorenanalyse der Kategorien ergab die Dimensionen
> - Fördernde/hemmende Einflüsse auf die Schulleistung
> - Angenehmer Schüler
> - Leistungsverhalten
> - Soziale Entwicklung,
>
> sowie vier Beurteilungstypen.
> In einem zweiten Durchgang beurteilten die gleichen Lehrer den leistungsstärksten und den leistungsschwächsten Schüler aus 35 Klassen anhand der Rating-Listen von HOFER (1969). Die Faktorenanalyse erbrachte die drei Dimensionen
> - Internale Bedingungen (Begabung/Anstrengung)
> - Soziales Verhalten
> - Seelische Robustheit,
>
> und einen starken ersten sowie einen schwachen zweiten Beurteilerfaktor.
> Die Untersucher führen die aus beiden Durchgängen resultierenden unterschiedlichen Faktoren auf die unterschiedlich determinierende Wirkung der beiden Erhebungsmethoden zurück.

In dieser Untersuchung wurde bei Benutzung des vorgegebenen Beschreibungsinstruments eine Einschränkung der Urteilsvariabilität gegenüber der freien Schülerbeschreibung

festgestellt. Der Unterschied in den resultierenden Faktoren bestand darin, daß bei vorgegebenen Ratinglisten internale Bedingungen (Begabung, Anstrengung) vorrangig betont wurden, während bei der freien Beurteilung äußere Einflüsse (häusliches Milieu, Förderungsmaßnahmen) in den Vordergrund traten, internale Bedingungen jedoch nur Komponenten anderer Dimensionen darstellten. Zu dieser Untersuchung sind allerdings Vorbehalte zu machen:

1. HUBER u. MANDL (1979) extrahierten aus den freien Beurteilungen vier Faktoren, aus den Beurteilungen per Ratingbogen dagegen nur drei Faktoren. Aus der Extraktionsposition der Faktoren wird geschlossen, daß "die Lehrer bei freier Beurteilung vor allem äußere Einflüsse (häusl.Milieu, Förderungsmaßnahmen, usw.) betonen", (a.a.O.,S.9), der Ratingbogen dagegen Urteilsbereiche akzentuiere, die im psychologischen Beurteilungsraum der Lehrer dieser Stichprobe untergeordnete Bedeutung haben, d.h. keinen eigenen Faktor bilden. Dieser Argumentation ist die Beobachtung aus der vorliegenden Untersuchung entgegenzuhalten, daß sich bei der Extraktion unterschiedlicher Faktorenanzahlen die Extraktionsränge und die Inhalte der Faktoren grundlegend ändern können, wie wir am Vergleich der Eigenwertlösung und der 4-Faktoren-Lösung gesehen haben. Zu mehreren Zeitpunkten wurde festgestellt, daß sich Variablen, die sich in der größeren Faktorenlösung als Komponenten verschiedener Faktoren zeigen, in einer kleineren Faktorenlösung sehr wohl zu einem eigenen Faktor sammeln können (z.B. Faktor IV: 'Robustheit ' der 4-Faktoren-Lösung zum Meßzeitpunkt 1; Faktor IV: 'emotionale Bewertung' zum Meßzeitpunkt 4, etc.).
Letztlich besagt der Vergleich der Faktoren aus den nach unterschiedlichen Erhebungsmethoden gewonnenen Urteilen, daß man eben das als Faktoren aus den Urteilen herausholen kann, was man vorher in Form der aus den verschiedenen Erhebungsmethoden resultierenden unterschiedlichen Beurteilungskriterien hineingesteckt hat, wobei man allerdings der Methode der freien Schülerbeschreibung eine größere Lebensnähe für den ökologischen Raum zubilligt, in dem diese Beurteilungen tagtäglich durchgeführt werden.

2. Im geschilderten Vergleich von HUBER u. MANDL (1979) zwischen den beiden Erhebungsmethoden scheint ein gewisser systematischer Einfluß die Vergleichbarkeit beider Vorgehensweisen zu beeinträchtigen; es erscheint fraglich, ob die aus beiden Methoden resultierenden unterschiedlichen Faktoren ausschließlich auf die verwendeten Methoden zurückzuführen sind. Die Autoren lassen in der freien Beschreibung gute, mittlere und schlechte Schüler (nach Leistungsniveaus zufällig ausgesucht) beurteilen, während mit HOFERs Eigenschaften-Liste nur der leistungsstärkste und der leistungsschwächste Schüler der Klasse beurteilt werden soll. Diese Fokussierung ausschließlich auf den sehr guten und sehr schlechten Schulerfolg provoziert im zweiten Durchgang der Untersuchung eine Kausalattribuierung, die die Ursache der Schulleistung im Schüler selbst lokalisiert in Form einer Zuschreibung größerer oder geringerer Fähigkeiten, Anstrengungsbereitschaft, etc.

   Im ersten Untersuchungsgang werden dagegen keine extremen Reizobjekte beurteilt, sondern ein Klassendurchschnitt aus leistungsstarken, mittleren und schwachen Schülern. Hierbei wird eine andere Kausalattribuierung aktiviert, die die Begründung für schulische Erfolge und Mißerfolge in den fördernden und hemmenden Einflüssen auf die Schulleistung sucht (häusl. Milieu, Förderungsmaßnahmen,etc.). Einen solchen Faktor "häusliches Milieu" haben wir in unserer Faktorenanalyse der drei Schülerstereotypen ebenfalls erhalten, allerdings nicht an so vorrangiger Stelle. Dieser Faktor war jedoch trotz gleichbleibender Eigenschafts-Listen im weiteren Verlauf der Beurteilungen nur noch selten zu reproduzieren, weswegen wir ihn auf die spezielle Themenstellung bei der Stereotypenbeurteilung zurückführten und ihn als Methodenfaktor ansahen.

3. Die Faktorenanzahl aus den freien Schülerbeschreibungen ist weniger der größeren Differenziertheit der Lehrerurteile als eher der Kategorisierungsleistung der Autoren in der Inhaltsanalyse zuzuschreiben, die ihre 274 Beschreibungseinheiten eben gerade in 36 Beurteilungskategorien unterbrachten, HOFER dagegen in nur 25 Kategorien. Die hohe Korrelation zwischen der Anzahl der Variablen und der Anzahl der resultierenden Faktoren hat HOFER selbst aufgezeigt.

4. Leider haben die Autoren keinerlei Hinweise bezüglich der statistisch-methodischen Bearbeitung der Daten aus den freien Schülerbeurteilungen angegeben. Es handelt sich hier wohl um den recht schwierigen Fall ein- oder mehrklassiger qualitativer Variablen, die nicht mit gleicher Nennungshäufigkeit auftreten und zudem noch von Vp zu Vp eine andere, urteilerspezifische Kombination bilden. So ist es auch müßig, Mutmaßungen darüber anzustellen, ob nicht etwa die vier resultierenden Beurteilungstypen durch ungleich häufiges Auftreten einzelner Variablen bzw. Variablengruppen in den Nennungen der Vpn schon im Datenmaterial angegeben waren und folglich notwendigerweise im faktoriellen Resultat der Q-sort Analyse wieder auftreten mussten.

Als Schlußfolgerung aus dieser Untersuchung sollte jedoch in künftigen Untersuchungen zur IPT eine lebensnähere Erhebung der Daten und eine der Eigenart der Daten gerechter werdende statistische Methodik angestrebt werden (z.B. Abgehen von der Implikation linearer Beziehungen zwischen den Variablen), für die GIGERENZER u. STRUBE (1978, S.84 f) den Vorschlag einer 'Nonmetrischen Multidimensionalen Skalierung' machen, die es gestattet, Fragen der Metrik und der Dimensionalität empirisch zu entscheiden.

VII Zusammenfassung

Die vorliegende Studie versuchte, folgende Fragen zu beantworten :

1.) Existieren in Lehrerurteilen Eigenschaftszusammenhänge, die man als die 'implizite Persönlichkeitstheorie' dieser Lehrer ansehen kann?

2.) Wieviele und welche Dimensionen weist die 'implizite Persönlichkeitstheorie' der untersuchten Lehrerstichprobe auf?

3.) Sind Zusammenhänge zwischen Skalen der Eigenschaftsbeurteilungen von Schülern und Persönlichkeitsmerkmalen von Lehrern festzustellen?

4.) Beeinflußt die 'implizite Persönlichkeitstheorie' der Lehrer die Struktur aktueller Schülerbeurteilungen?

5.) Verändert sich die Struktur aktueller Urteile mit zunehmender Bekanntschaft von Beurteilern und Beurteilten?

6.) Welche Dimensionen der Urteilsstrukturen bleiben über die Zeit hinweg gleich, welche verändern sich, treten im Kontaktzeitraum neue Dimensionen hinzu, verschwinden bereits etablierte Dimensionen?

7.) Ist in den Urteilen über die Meßzeitpunkte ein bestimmtes Bezugsschema festzustellen?

8.) Wie stabil sind die Urteile über Schülereigenschaften im Untersuchungszeitraum?

Eine Stichprobe von 23 Lehrern der ersten Klassen der zwei Gymnasien eines Ortes wurden nach den Merkmalen befragt, nach denen sie gewöhnlich die Persönlichkeit ihrer Schüler beurteilen; zu den in Form von Eigenschaftsbegriffen vorliegenden Merkmalen wurden von den Lehrern die von ihnen benutzten sprachlichen Gegenteile erfragt, wobei sich in der weitgehenden Deckungsgleichheit der genannten Gegenteile eine weitgehende Einheitlichkeit der Verwendung der Begriffe abzeichnete.
Die Eigenschaftsbezeichnungen und ihre sprachlichen Gegenteile wurden nach Häufigkeitsauszählung und Eliminieren der seltensten Nennungen zu 29 bipolaren, siebenstufigen Rating-Skalen angeordnet und bildeten die zukünftig verwendete Beurteilungsliste.

1.) Es waren Zusammenhänge der Eigenschaftsbeurteilungen festzustellen, die nach Maßgabe der verschiedenen Untersuchungsmethoden teilweise auch in anderen Untersuchungen festgestellt worden waren.

2.) Je nach Extraktionskriterium wurden bis zu sieben Faktoren festgestellt, wobei jedoch durchgängig nur drei Faktoren signifikant waren, die wie folgt benannt wurden: I. Geistige Regsamkeit (wesentliche Ladungen auf den Eigenschaften 'eigenes Mitdenken', 'intelligent', 'begabt', 'kreativ', 'aufgeweckt', 'ehrgeizig', 'gute Mitarbeit', 'unterfordert'.)
II. Schuldisziplin (Ladungen auf: 'fleißig', 'zuverlässig', 'ordentlich', 'konzentriert', 'ausdauernd', 'stört selten', 'positive Einstellung zu Schule und Lernen', 'höflich', 'ehrlich'.)
III. Kontaktoffenheit (Ladungen auf: 'kontaktfreudig', 'guter Lehrerkontakt', 'offener Charakter', 'hübsch'.)

3.) Es konnten keine Zusammenhänge von Testwerten eines Persönlichkeitstests (FPI) und Skalen der Beurteilung von Schülerstereotypen festgestellt werden.

4.) Die signifikanten Faktoren aus aktuellen Beurteilungen nach kurzer Bekanntschaft zwischen Lehrern und Schülern waren hochkongruent mit den Faktoren der 'impliziten Persönlichkeitstheorie' der Lehrer; dieses Ergebnis wird als Deckungsgleichheit der Ähnlichkeitsschätzungen vorgegebener Eigenschaftsbegriffe interpretiert, die der IPT und den Urteilen in einer informationsreduzierten Situation zugrundeliegen.

5.) Ein spezifischer Faktor 'häusliches Milieu', der in der Faktorenanalyse der leistungsbetonenden Schülerstereotypen extrahiert werden konnte, in der Analyse der aktuellen Beurteilungen relativ unbekannter Schüler jedoch nicht, läßt an einer Ähnlichkeitsschätzung von Eigenschaften als ausschließlicher Quelle der IPT zweifeln; die jeweils gestellte Beurteilungsaufgabe scheint die Art der resultierenden Faktoren wenigstens bis zu einem gewissen Grad zu beeinflussen. Die inkongruenten Faktoren der IPT und der Faktorenstruktur der aktuellen Beurteilungen waren jedoch methodisch kaum zu vertreten, weswegen sie in dieser Diskussion nicht als Argumente verwendet werden können.

6.) In der standardmäßig durchgeführten Eigenwert-Lösung zeigte sich eine allgemeine methodische Schwäche der letzten zwei oder drei Faktoren, woraus geschlossen wurde, daß die psychologisch "richtige" Zahl der Faktoren unter der mathematisch "richtigen" Zahl lag; in der Gegenüberstellung der einzelnen Faktorenlösungen nach Kriterien von FÜRNTRATT (1969) stellte sich die 4-Faktoren-Lösung als best-interpretierbare heraus.

7.) Bis auf einige Ausnahmen liessen sich in der Eigenwert- und in der 4-Faktoren-Lösung alle Faktoren des Meßzeitpunktes 1 (relativ unbekannte Schüler) zu jedem weiteren Meßzeitpunkt mit steigendem Bekanntheitsgrad zwischen Urteilern und Beurteilten replizieren.

8.) Die Ausnahmen von Punkt 8 betrafen eine Kontamination der Inhalte des ersten und zweiten extrahierten Faktors zu den Zeitpunkten 4,5,10 und 11 zu einem bisher nicht existenten Faktor "Schuleignung". Da die Untersuchungsbedingungen und die statistische Methodik zu diesen Zeitpunkten keine Besonderheiten gegenüber den anderen Zeitpunkten aufwiesen, wird dieses in allen Faktorenlösungen auftretende Phänomen der Besonderheit der Urteilssituation zugeschrieben: Zu den genannten Zeitpunkten hatten die urteilenden Lehrer Entscheidungen über den weiteren Verbleib der Schüler auf dem Gymnasium zu treffen. Aufgrund dieser speziellen Beurteilungsaufgabe wird ein Abweichen der Faktoren aus aktuellen Beurteilungen von den Faktoren der IPT für möglich gehalten.

9.) Der Vergleich der verschiedenen Faktorenlösungen bezüglich der zentralen Faktoren ergab ein uneinheitliches Bild: je nach Faktorenlösung und Meßzeitpunkt wechselten die Faktoren, die den größten Varianzbetrag banden. Ein inhaltlich gleicher zentraler Faktor kann in dieser Arbeit also nicht als Argument besonderer Wichtigkeit einer Dimension für die Beurteiler verwendet werden; er ist vielmehr von der gewählten Faktorenlösung abhängig.

10.) Es konnte keine durchgängige 'zentrale Eigenschaft' in den Analysen festgestellt werden.

11.) Die Untersuchung der Stereotypen 'der gute Schüler', 'der schlechte Schüler' und 'der Schüler der 1. Gymnasialklasse' zeigte ein einheitlich positiv gefärbtes Bild vom 'guten' Schüler. Demgegenüber war das Bild vom 'schlechten Schüler' nicht einheitlich negativ. Auch der schlechte Gymnasiast der 1. Klasse wird durchschnittlich als intelligent und begabt angesehen. Hier stellte sich allerdings eine Dichotomisierung der Bilder vom schlechten Schüler in der Stichprobe heraus: Der geringe Schulerfolg wird einerseits auf die Kombination der Eigenschaften 'intelligent/faul' und andererseits 'unintelligent/fleißig' zurückgeführt.

12.) Ein einheitliches Bezugssystem der Urteile im Sinne der Strenge oder Milde der Beurteilung konnte nicht festgestellt werden. Hohe oder geringe Mittelwerte einzelner Skalengruppen bezogen sich weniger auf einzelne Dimensionen des Urteils als eher auf die von den Lehrern geschätzte Relevanz der Skalen für schulisches Beurteilen. Besonders hohe Mittelwerte mit relativ geringen Streuungsbreiten über den Meßzeitraum wurden bei als relativ irrelevant eingeschätzten Skalen mit schulmoralisch wertenden Eigenschaftsbezeichnungen festgestellt (ehrlich, höflich, sympathisch, positive Einstellung zu Schule und Lernen) während auf den Skalen mit spezifischer Relevanz für nur einige Lehrer besonders geringe Mittelwerte für 'sensibel-', 'überfordert-', 'akzeleriert-' festgestellt wurden, die eventuell auf eine geringere konstitutionelle Belastbarkeit, ein 'asthenisches Syndrom' hinweisen.

13.) Die Stabilität der Urteile über den Meßzeitraum nahm rhythmisch progredient zu; die Koeffizienten erreichten jedoch auch nach einem halben Schuljahr noch nicht die Höhe, daß mit einiger Berechtigung von einem feststehenden, sich kaum noch verändernden 'image' der Schüler gesprochen werden kann. Erstaunlich und aus der Erhebungssituation nicht zu erklären ist eine geringe Stabilität der Urteile nach einem zeitlichen Intervall von 6 Monaten zwischen den letzten beiden Meßzeitpunkten, die auch Variablen umfaßte, die sich mit Sicherheit in diesem Intervall nicht gravierend veränderten (z.B. häusliches Milieu).

14.) In einer abschließenden - die Methodik dieser und ähnlicher Untersuchungen reflektierenden - Kritik wurde festgestellt, daß viele erhebungs- und statistisch-methodische Implikationen von im Forschungszusammenhang der 'impliziten Persönlichkeitstheorien' üblichen Verfahren die Validität der Ergebnisse für den gemeinten Sachverhalt des Zusammenhangs zugeschriebener Eigenschaften in der Alltags-Diagnostik von Lehrern empfindlich senken.

Überspitzt formuliert haben die vorliegende und andere Untersuchungen nur eine inhaltliche Validität für Situationen, in denen Lehrer anhand vorgegebener Eigenschaftsinventare (z.B. Schülerbeurteilungsbogen) explizit Urteile formulieren.

Als Schlußfolgerung aus den Erfahrungen, die der Autor im Verlauf der Arbeit machte, wird die Empfehlung für weitere Arbeiten zum Themenkreis der IPTn gezogen, lebensnähere Erhebungsmethoden und datengerechtere statistisch-mathematische Modelle als die Faktorenanalyse zu verwenden, um mit größerem Anspruch auf Validität der Untersuchungsergebnisse das Phänomen naiver Dispositionstheorien zu erfassen.

Anderson, N.H.: Application of a linear-serial model to a personality-impression task using serial presentation.
J.Pers.Soc.Psychol.,10(4), 1968, 354-62

Asch, S.E.: Forming impressions of personality.
J.abn.soc.Psychol., 41, 1946, 258-290

Asch, S.E.: Studies of independence and conformity: a minority of one against an unanimous majority.
Psychol.Monogr.,1956,70,No.9,Whole No.416.

Babbie, E.R.: Survey Research Methods.
Belmont, 1973. Zit.nach Helmreich, R.,S.11 f.

Bach, T.R.: Adjustment differences related to pattern of rating of the other.
Psychol.Rep., 32(1), 1973, 19-22

Bastine, R.: Fragebogen zur direktiven Einstellung.
Göttingen, 1971.

Baumann, Urs.: Psychologische Taxometrie. Eine Methodenstudie über Ähnlichkeitskoeffizienten, Q'-Clusteranalyse, Q^F-Faktorenanalyse.
Bern, 1971.

Beach, C.,Wertheimer, K.: A free response approach to the study of person cognition.
J.abn.soc.Psychol., 62, 1961, 367-374.

Bernhardson, C.S., Fisher, R.J.: Perceptions of social desirability and frequency of occurence of traits and the probability that the traits will be endorsed.
Percept.mot.skills, 30(2), 1970, 571-576.

Beutel, P.et.al.: SPSS - Statistik-Programmsystem für die Sozialwissenschaften
Stuttgart, 1976.

Beutel, P.et.al.: SPSS 7 - Statistik-Programmsystem für die Sozialwissenschaften.
Stuttgart, 1978[2].

Bierhoff, H.W., Bierhoff-Alfermann, D.: Interindividual differences in the application of implicit personality theories: correlations between cognitive structure and personality in high school students (10th grade).
Arch.Psychol.,127(3-4), 1975, 284-299.

Bierhoff, H.W., Bierhoff-Alfermann, D.: The use of the psychological theories by naive judges: a study in implicit personality theory.
Europ.J.Soc.Psychol.,6(4), 1976, 429-445.

Bieri, J.: Cognitive complexity-simplicity and predictive behaviour.
J.abn.soc.Psychol., 51, 1955, 263-268.

Birnbaum, M.H.: The nonadditivity of personality impressions.
J.exp.Psychol., 102, 1974, 543 - 561.

Bleda, P.R.: Toward a clarification of the role of cognitive and affective processes in the similarity-attraction relationship.
J.Pers.Soc.Psychol., 29, 1974, 368-373.

Bogen, H.: Psychologische Grundlegung der prakt.Berufsberatung.
Langensalza, 1927. Zit.nach Hasemann, K.: Verhaltensbeobachtung. in: Hb.d.Psycholigie, Bd 6, Psycholog.Diagnostik.

Borman, W.C., Graham, W.K.: Polarity and accuracy of ratings and the meaningfulness of personality dimensions.
H.Psychol., 84, 1973, 55 - 60.

Bourne, E.: Can we desribe an individual's personality?-Agreement on stereotypes vs. individualattributes.
J..Pers.Soc.Psychol., 35(12), 1977, 863-872.

Brickenkamp,R.(Hrsg.): Handbuch psychologischer und pädagogischer Tests.
Göttingen, 1975.

Brigham, H.C.: Ethnic stereotypes and attitudes: a deferrent mode of analysis.
J.Pers., 41(2), 1973, 206-223.

Bruner, J.S.: Personality dynamics and the process of perceiving.
In: Blake, R., Ramsey, G.V., (Hrsg.): Perception - an approach to personality. N.Y. 1951.

Bruner, J.S., Tagiuri, R.: The perception of people.
In: Lindzey, G.(Hrsg.): Handbook of social psychology, Vol. 2, 1954, 601-633.

Cantor, J.H.: Individual needs and salient constructs in interpersonal perception.
J.Pers.Soc.Psychol., 34(3), 1976, 519-525.

Cattell, R.B., Coan, R.W.: Child personality structure as revealed in teachers behaviour ratings.
J.clin.Psychol., 13, 1957, 315-327.

Cattell, R.B.: Personality and motivation structure and measurement.
N.Y., 1957.

Cattell, R.B., Nesselrode, J.: Untersuchung der interkulturellen Konstanz der Persönlichkeitsfaktoren im 16-PF-Test.
Psychol.Beiträge, 8, 1965, 502-515.

Chaikin, A., Derlega, V.J.: Liking for the norm-breaker in self-disclosure.
J.Pers., 42(1), 1974, 117-129.

Cline, V.B.: Interpersonal perception.
In: Maher, B.A. (Hrsg.): Progress in experimental personality research.
Vol.1,1964,221-284.

Cohen, R.: Systematische Tendenzen bei Persönlichkeitsbeurteilungen.
1969.

Cohen, R., Schümer, R.: Eine Untersuchung zur sozialen Urteilsbildung.
I. Die Verarbeitung von Informationen unterschiedlicher Konsonanz.
Unveröff.Manuskript, Hamburg, 1967.
Zit.nach Cohen, 1969, S.61.

Costin, F., Grush, J.E.: Personality correlates of teacher-student behavior in the college classroom.
J.educ.Psychol., 65(1), 1973, 35-44

Crandall, J.E.: Predictive value and confirmability of traits as determinants of judged trait importance.
J.Pers., 38(1), 1970, 77-90.

Cronbach,L.J. : Processes affecting scores on ''understanding of others'' and assumed similarity''.
Ps.Bull. 52,1955,177-193.

Cronbach, L.J.: Proposals leading to analytic treatment of social perception scores.
In: Tagiuri, R., Petrullo, L. (Hrsg.): Person perception and interpersonal behavior.
Stanford, 1958, 353-379.

Cronbach, L.J.: Essentials of psychological testing.
N.Y., 1960[2] .

Crowther, B.: Differentiating ability in the use of personality scales.
Percept.mot.skills, 33(1), 1971, 325-326.

Deaux, K., Farris, E.: Complexity, extremity, and affect in male and female judgments.
J.Pers., 43(3), 1975, 379-389.

Delia, J.G., Gonyea, A.H., Crockett, W.H.: The effects of subject-generated and normative constructs upon the formation of impressions.
Br.J.Soc.Clin.Psychol., 10(4), 1971, 301-305.

Diggins, D.: The role of social and nonsocial traits in interpersonal attraction.
J.Pers., 42(3), 1974, 345-359.

Dion,K., Berscheid,E.,Walster,E.: What is beautiful is good.
J.Pers.Soc.Psychol., 24(3), 1972, 285-290.

Donat, H.: Persönlichkeitsbeurteilung. Methoden und Probleme der Charaktererfassung im pädagogischen Bereich. München, 1965.

Dornbusch, S.M., Hastorf, A.H., Richardson, S.A., Muzzy, R.E., Vreeland, R.S.: The perceiver and the perceived: Their relative influence on the categories of interpersonal cognition. J.Pers.+Soc.Psychol., 1, 1965, 434-440.

Dorsch, F.: Psychologisches Wörterbuch. Bern, 1963[7].

Drewery, J.: An interpersonal perception technique. Br.J.Med.Psychol., 42(2), 1969, 171-181.

Eagle, M., Goldberger, L., Breitman, M.: Field dependence and memory for social vs neutral and relevant vs irrelevant incidental stimuli. Percept.mot.skills, 29(3), 1969, 903-910.

Ehrlich, H.J.: Affective style as a variable in person perception. J.Pers., 37(4), 1969, 522-540.

Ekehammar, B., Magnusson, D.: Intertrait correlations as a function of amount of judgmental information about the ratee. Percept.mot.skills, 36(2), 1973, 373-374.

Emmerich, W.: Development trends in evaluations of single traits. Child Dev., 45(1) 1974, 172-183.

Engelmayer, O.: Beobachtung und Beurteilung des Schulkindes. 1949.

Engelmayer, O.: Psychologie für den schulischen Alltag. Das psychologische Grundwissen im Dienste einer modernen Unterrichtspraxis. München, 1960.

Eyferth, K., Sixtl. F.: Bemerkungen zu einem Verfahren zur maximalen Annäherung zweier Faktorenstrukturen aneinander. Arch.ges.Ps. 117, 1965, 131-138.

Fahrenberg, J., Selg, H.: Das Freiburger Persönlichkeitsinventar FPI. Göttingen, 1970.

Feather, N.T., Simon, J.G.: Reactions to male and female success and failure in sex-linked occupations; impressions of personality, causal attributions, and perceived likelihood of different consequences. J.Pers.Soc.Psychol., 31(1), 1975, 20-31.

Feigel, W., Keitel, E.: Bayerische Landesvolksschulordnung. München, 1964[3]. Zit.nach: Ulich, D. und Mertens, W.: Urteile über Schüler. Weinheim, 1973.

Ferguson, L.W.: Spearman-Brown reliability coefficients for Ferguson's 80-item pair-comparisons person-perception scale. Percept.mot.skills, 38(3), 1974, 1161-1162.

Feshbach, N.D.: Student teacher preferences for elementary school pupils varying in personality characteristics.
J.Educ.Psychol., 60(2), 1969, 126-132.

Fischer, G., Roppert, J.: Bemerkungen zu einem Verfahren der Transformationsanalyse.
Arch.Ges.Psychol., 116, 1965, 98-100.

Friedland, S.J., Crockett, W.H., Laird, J.D.: The effects of role and sex on the perception of others.
J.Soc.Psychol., 91(2), 1973, 273-283.

Fürntratt, E.: Zur Bestimmung der Anzahl interpretierbarer gemeinsamer Faktoren in Faktorenanalysen psychologischer Daten.
Diagnostica, 15, 1969, 62-75.

Gormly, J., Edelberg, W.: Validity in personality trait attribution.
Am.Psychol., 29(3), 1974, 189-193.

Goodenough, F.L.: Draw-a-man-Test.
N.Y., 1926.

Graumann, C.F.: 'Social perception'. Die Motivation der Wahrnehmung in neueren amerikanischen Untersuchungen.
Z.exp.angew.Psychol., 3, 1956, 605-661.

Graumann, C.F.: Nicht-sinnliche Bedingungen des Wahrnehmens.
In: W.Metzger (Hrsg.): Allgemeine Psychologie. Der Aufbau des Erkennens. 1.Hb.: Wahrnehmung und Bewußtsein.
Hb.d.Ps.Bd.1.Göttingen, 1966, 1031-1096.

Graumann, C.F.: Sozialpsychologie: Ort, Gegenstand und Aufgabe.
In: Ders:(Hrsg.) Hb.d.Ps.Bd.VII/1 (Sozialpsychologie).
Göttingen, 1969, 30.

Gruenfeld, L., Arbuthnot, J.: Field Independence as a conceptual framework for prediction of variability in ratings of others.
Percept.mot.skills, 28(1), 1969, 31-44.

Grund-und Strukturdaten: Der Bundesminister für Bildung und Wissenschaft, 1976, 28.

Guilford, J.P.: Psychometric methods.
N.Y., 1954^2.

Gulliksen, H.(Hrsg.): Contributions to mathematical psychology.
Hole, 1964. Hier: Tucker, L.R.: The extension of factor analysis to three dimensional matrices.

Grush, J.E., Clore, G.L., Custin, F.: Dissimilarity and attraction: when difference makes a difference.
J.Pers.Soc.Psychol., 32(5), 1975, 783-789.

Guttman, L.: Some necessary conditions for common-factor analysis.
Psychometrika, 19, 1954, 149-161.
Zit.nach: Pawlik, K.: Dimensionen des Verhaltens.

Hamilton, D.L.: The structure of personality judgments: comments on Kuusinen's paper and further evidence.
Scand.J.Psychol., 11(4), 1970, 261-265.

Hampe, E.: Parents and teachers perceptions of personality characteristics of children selected for classes for the learning disabled.
Psychol. Rep., 37(1), 1975, 183-189.

Hanno, M.S., Jones, L.E.: Effects of a change in reference person on the multidimensional structure and evaluations of trait adjectives.
J.Pers.Soc.Psychol., 28(3), 1973, 368-375.

Harman, H.: Modern factoranalysis.
Chic.,1960.

Harris, E.W. (Hrsg.):Problems in measuring change.
Univ.of Wisconsin Press, Madison, 1963.

Hartig, M., Petermann, F.: Probleme und Methoden der Psychotherapieforschung.
1975.

Hastorf, A.H., Schneider, D.J., Polefka, J.: Person Perception.
1970.

Hays, W.L.: An approach to the study of trait implication and trait similarity.
In: Tagiuri, R., Petrullo, L. (eds.): Person perception and interpersonal behavior.
Stanford, 1958, 289-299.

Helmreich, R.: Strategien zur Auswertung von Längsschnittdaten.
Stuttgart, 1977.

Hendrick, C.: Preference for consistent and inconsistent information in impression formation.
Percept.mot.skills, 28(3), 1969, 877-878.

Hendrick, S.: Effects of salience of stimulus inconsistency on impression formation.
J.Pers.Soc.Psychol., 22(2), 1972, 219-22.

Himmelfarb, S.: Integration and attribution theories in personality impression formation.
J.Pers.Soc.Psychol., 23(3), 1972, 309-313.

Höhn, E.: Der schlechte Schüler. Sozialpsychol.Untersuchungen über das Bild d. Schulversagers.
München, Piper, 1967.

Hörmann, H.: Theoretische Grundlagen des projektiven Tests.
In: Heiss, R. et.al. (Hrsg.): Hb.d.Ps., BD 6: Psychol. Diagnostik.
Göttingen, 1964, 71-112.

Hörmann, H.: Psychologie der Sprache.
Berlin, 1967.

Hofer, M.: Die Schülerpersönlichkeit im Urteil des Lehrers.
Weinheim, 1969.

Hofstätter, P.R.: Gruppendynamik.
Hamburg, 1957.

Hofstätter, P.R.: Einführung in die Sozialpsychologie.
Kröner, Stuttgart, 1963.

Hofstätter, P.,Wendt, D.: Quantitative Methoden der Psychologie.
Frankfurt, 1974[4].

Hohlen, M.: SPSS - 6000 Version 7.0 Update.
Vervielfältigtes Manual No. 437 der Northwestern University
Vogelback Computing Center, Evanston, 1977.

Hollingworth, H.L.: Experimental studies in judgment.
N.Y. 1910. Zit.nach: Guilford, J.P.: Psychometric methods,
1954[2], 323.

Hollingworth, H.L.: The central tendency of judgment.
J.Phil.Psychol., sci.Meth.,7,1910, 461-468.
Zit.nach: Guilford, J.P.: Psychometric methods.
1954[2], 323.

Horn, W.: Begabungstestsystem.
Göttingen, 1956. (Zit.S.4).

Huth, A.: Meine Schüler - Eine Beobachtungsanleitung für den Lehrer.
Ansbach, 1955[3].

Jasper, Lawrence, G.:Implicit personality theory: reliability of measurement.
Dissertation abstracts international, 35(10-B), 1975, 5114-5115.

Jellison, J.M., Zeisset, P.T.: Attraction as a function of the commonality and
desirability of a trait shared with another.
J.Pers.Soc.Psychol.,11(2), 1969, 115-120.

Jones, E.E., DeCharms, R.: The organizing function of interaction roles in
person perception.
J.abn.Soc.Psychol., 57, 1958, 155-164.

Jones, S.C., Knurek, D.A., Regan, D.T.: Variables affecting reactions to
social acceptance and rejection.
J.Soc.Psychol., 90(2), 1973, 269-284.

Kaltsounis, B., Higdon, G.: Black teacher's perceptions of the ideal pupil.
Percept.mot.skills, 45 (3PT2), 1977, 1071-1075.

Kaminski, G.: Das Bild vom Anderen.
Berlin, 1959, 136.

Kaminski, G.: Zur Analyse von Unrteilssequenzen.
Ps.+Praxis, a,3,1959, 184-192.

Kaminski, G.: Die Beurteilung unserer Mitmenschen als Prozeß.
In Lienert, G.(Hrsg.): Ber.23.Kongreß DGfP.
Hogrefe, Göttingen, 1963,51-67.

Kamis, E.G.: Situational determinants of implicit personality theories. Dissertation abstracts international, Vol.34(6-8), 1973, 2900-2901.

Kaplan, M.F.: Context effects in impression formation: the weighted average versus the meaning-change formulation. J.Pers.Soc.Psychol., 19(1), 1971, 92-9.

Kaplan, M.F.: Stimulus inconsistency and response dispositions in forming judgments of other persons. J.Pers.Soc.Psychol., 25(1), 1973, 58-64.

Kaplan, M.F.: Context-induced shifts in personality trait evaluation: a comment on the evaluative halo effect and the meaning change interpretations. Psychol.Bull.,81(11), 1974, 891-895.

Kaplan, M.F.: Measurement and generality of response dispositions in person perception. J.Pers., 44(2), 1976, 179-194.

Kelly, H.H.: The warm-cold variable in first impressions of persons. J.Pers., 18, 1950, 431-439.

Kelly, G.A.: The psychology of personal constructs. Vol.1, N.Y. 1955.

Kieferle, D.A., Sechrest, L.: Effects in alterations in personal constructs. J.Psychol.Studies, 12, 1961, 173-178.

Koltuv, B.B.: Some characteristics of intrajudge trait intercorrelations. Psychol.Monogr.76, No.552, 1962.

Krebs, D., Adinolf, A.A.: Physical attractiveness, social relations, and personal style. J.Pers.Soc.Psychol., 31(2), 1975, 245-253.

Kuusinen, J., Nystedt, L.: The convergent validity of four indices of cognitive complexity in person perception, a multi-index multimethod and factor analytical approach. Scand.J.Psychol., 16(2), 1975, 131-136.

Kuusinen, J., Nystedt, L.: Individual versus provided constructs, cognitive complexity and extremity of ratings in person perception. Scand.J.Psychol., 16(2), 1975, 137-142.

Labbie, S.: Primacy-regency effects in impression formation and congruity-incongruity of stimulus material. Percept.mot.skills, 37(1), 1973, 275-278.

Labouvie, G.V., Baltes, P.B.: Adolescent perception of adolescent change in personality and intelligence. J.Genet.Psychol., 122(2), 1973, 269-285.

Lay, C.H., Jackson, D.N.: Analysis of the generality of trait-inferential relationships.
J.Pers.Soc.Psychol., 12(1), 1969, 12-21

Lersch, P: Aufbau der Person.
Barth, Mü, 1954.

Levenson, H., Mahler, I.: Attitudes toward others and components of internal-external locus of control.
Psychol.Rep., 36(1), 1975, 209-210.

Leventhal, H.: The effects of set and discrepancy on impression change.
J.Pers.30, 1962, 1-15.

Levin, I.P., Wall, L.L., Dolezal, J.M., Norman, K.L.: Differential weighting of positive and negative traits in impression formation as a function of prior exposure.
J.Exp.Psychol., 97(1), 1973, 114-115.

Levy, L.H., Dugan, R.D.: A constant error approach to the study of dimensions of social perception.
In: J.of.abn.Soc.Psychol., 61, 1960, 21-24.

Luchins, A.S., Luchins, E.H.: Effects of preconceptions and communications on impressions of a person.
J.Soc.Psychol., 81(2), 1970, 245-252.

Maguire, T.V.: Implicit theories of personality: the influence of role - relations.
Dissertations Abstracts International, 35(7-8), 1975, 3588.

McCall, J., Rae, G.: Inferential sets and person perception: a partial replication and extension.
Percept.mot.skills, 37(2), 1973, 479-482.

Merz, F.: Über die Beurteilung der persönl. Eigenart unter Mitmenschen.
Arch.f.d.ges.Psychol., 114, 1962, 187-211.

Miller, A.G., Gillen, B., Schenker, C., Radlove, S.: The prediction and perception of obedience to authority.
J.Pers. 42(1), 1974, 23-42.

Mirels, H.L.: Implicit personality theory and inferential illusions.
J.Pers., 44(3), 1976, 467-487.

Newcomb, T.: An experiment designed to test the validity of a rating technique.
J.Educ.Ps., 22, 1931, 279-289.

Nie, N.H., Hull, C.H., Jenkins, J.G., Steinbrenner, K., Bent, D.H.: SPSS-Statistical package for the social sciences.
N.Y., 1975^2.

Nötzold, J.: Hinweisung zur Bearbeitung statistischer Probleme an der CYBER 175.
Unveröffentl. Manuskript des Rechenzentrums d.Rheinl.-Westf.-Technischen Hochschule Aachen, 1977.

Nystedt, L., Kuusinen, J., Ekehammar, B.O.: Structural representations of person perception: a comparison between own and provided constructs.
Scand.J.Psychol., 17(3), 1976, 223-233.

Oltman, P.K., Goodenough, D.R., Witkin, H.A., Freedman, N., Friedman, F.: Psychological differentiation as a factor in conflict resolution.
J.Pers.Soc.Psychol., 32(4), 1975, 730-736.

Osgood, C.E.: The nature and measurement of meaning.
Psychol.Bull. 49, 1952.

Osgood, C.E.: Psycholinguistics.
In: Koch, S. (Hrsg.): Psychology: a study of a science, Bd VI.
N.Y., 1963, 244-316.

Padgitt, S.T.: Implicit personality theory: the effect of sexual responses on evaluative judgements.
Dissertation abstracts international, 35(6-8), 1974, 3.071.

Passini, F.T., Norman, W.T.: A universal conception of personality structure?
J.Pers.Soc.Psychol., 4, 1966, 44-49.

Perry, R.P.: The effect of selective attention on communicating personality impressions.
J.Soc.Psychol., 98(first half), 1976, 121-130.

Plüschke, R.: NOS Version 1 - Betriebssystem der CDC CYBER 175.
Vervielfältigtes Manuskript des Lehrstuhles für Unternehmensforschung an der Rheinl.-Westf.-Techn.-Hochschule, Aachen.

Reid, D.W., Ware, E.E.: Affective style and impression formation: reliability, validity and some inconsistencies.
J.Pers., 40(3), 1972, 436-450.

Revenstorff, D.: Über Profilähnlichkeit (Profile Similarity).
Arch.Psychol., 125(2), 1973, 203-232.

Rice, W.K. Jr.: Effects of discipline techniques on children's personality trait inferences.
J.Educ.Psychol., 67(4), 1975, 570-575.

Rinella, S., Ferguson, L.W., Sager, E.B.: Personality-impression formation as a function of visual cues and set.
Percept.mot.skills, 31(2), 1970, 427-430.

Rosenbaum, M.E.,Levin, I.P.: Impression formation as a function of source credibility and the polarity of information.
J.Pers.Soc.Psychol., 12(1), 1969, 34-37.

Rosenthal, R., Jacobson, L.: Pygmalion in the classroom. Teacher expectation and pupils' intellectual development.
N.Y., 1968.

Rosenthal, R.: Teacher expectations.
Persönliche Mitteilung durch Manuskript vom 20.3.1970.
Cambridge/Mass.

Rupp, H.: Über optische Analysen.
Psychol.Forschung 4, Göttingen 1955, 262-300.

Sarbin, T.R.: Role theory.
In: Lindzey, F.(ed.): Handbook of Social Psychology,
Cambridge, Mass. Addison-Wesley, 1, 1954, 223-258.

Scherer, S.E.: Influence of proximity and eye contact on impression formation.
Percept.mot.skills, 38(2), 1974, 538.

Schneider, D.J.: Implicit personality theory: a review.
Psychol.Bull., 79(91), 1973, 294-309.

Sechrest, L.: Personal constructs and personal characteristics.
J.Individ.Psychol., 24(2), 1968, 162-6.

Secord, P.F., Berscheid, E.S.: Stereotyping and the generality of implicit personality theory.
J.Pers.31, 1963, 65-78.

Shweder, R.A.: How relevant is an individual difference theory of personality?
J.Pers., 43, 1975, 455-484.

Shrauger, J.S., Patterson, M.B.: Self-evaluation and the selection of dimensions for evaluating tohers.
J.Pers., 42(4), 1974, 569-585.

Singh, R.: Attraction as a function of similarity in attitudes and personality characteristics.
J.Soc.Psychol., 91(1), 1975, 87-95.

Sleet, D.A.: Physique and social image.
Percept.mot.skills, 28(1), 1969, 295-299.

Slocombe, C.S.: Psychology of co-operation.
Personnel J., 16, 1938, 325-332.
Zit.nach: Hasemann, K.: Verhaltensbeobachtung. In: Hb.d. Psychol., Hogrefe, Bd.6, Psychol.Diagnostik, S.836.

Snodgrass, S.R.: The development of trait inference.
J.Genet.Psychol., 128(2), 1976, 163-72.

Stockford, L., Bissel, H.W.: Factors involved in establishing a merit rating scale.
Personnel, 26, 1949, 94-116.
Zit.nach: Cronbach, L.H.: Essentials of psychol.testing.
N.Y. 1960^2.

Todd, F.J., Rappoport, L.: A cognitive structure approach to person perception- a comparison of two models.
J.Abn.Soc.Psychol., 68, 1969, 469-478.

Tupes, E.C., Christal, R.E.: Stability of personality trait rating factors obtained under diverse conditions.
In: USAF WADC Technical Note, No.58-61, 1958.
Zit.nach: Ulich, D., Mertens, W.: Urteile über Schüler, Weinheim 1973, S.111.

Tzeng, O.C.: Differentiation of affective and denotative meaning systems and their influence in personality ratings.
J.Pers.Soc.Psychol., 32(6), 1975, 978-988.

Ulich, D., Mertens, W.: Urteile über Schüler.
Weinheim, 1973.

Van der Kloot, W.A.: The role of implicit personality theory in clarifying problems arising from studies of person perception.
Gedrag: Tijdschrift voor Psychologie, Vol.4(2), 1976, 80-97.

Walters, H.A., Jackson, D.N.: Group and individual regularities in trait inference: a multidimensional scaling analysis.
In: Multivariate behavior research, 1, 1966, 145-163.
Zit.nach: Hofer, M.: Die Schülerpersönlichkeit im Urteil des Lehrers,Weinheim, 1969.S.

Warr, P.B., Knapper, Ch.: The perception of people and events.
London, 1968.
Zit.nach: Graumann, C.F.: Sozialpsychologie: Ort, Gegenstand und Aufgabe. In: Ders.(Hrsg.): Hb.d.Psychol., Bd.VII/1 (Sozialpsychologie). Göttingen, 1969.

Weinschenk, C.: Das unmittelbare Gedächtnis als selbständige Funktion.
Göttingen, 1955.

Wewetzer, K.-H.: Intelligenztests für Kinder.
In: Heiss, R. (Hrsg.): Hb.d.Psychol., Bd. VI: Psychologische Diagnostik. Göttingen, 1964, S. 200-225.

Weissbach, T.A., Zagon, G.: The effect of deviant group membership upon impressions of personality.
J.Soc.Psychol., 95 (second half), 1975, 263-266.

Wyer, R.S. Jr.: Changes in meaning and halo effects in personality impression formation.
J.Pers.Soc.Psychol.,29(6),1974, 829-835.

Zajouc, R.B.: The process of cognitive tuning in communication.
J.Abn.Soc.Psychol., 61, 1960, 159-167.

Ziler, H.: Der Mann-Zeichentest.
Münster, 1958.

BECKMANN L.: Auswirkungen von schulischen Leistungen auf die Kausalattribuierung von lehrenden und beobachtenden Personen.
In: HOFER, M., WEINERT, F. (Hrsg.): Pädagogische Psychologie, Bd. 2. Lernen und Instruktion.
Frankfurt, Fischer 1973.

BRANDSTÄTTER, H.: Die Beurteilung von Mitarbeitern.
In: MAYER,A., HERWIG, B. (Hrsg.): Hdb. d. Psychol., IX, Betriebspsychologie, Göttingen, 1970^2, S.695.

BROPHY,J.;GOOD,T.: Die Lehrer-Schüler-Interaktion.
München: Urban + Schwarzenberg, 1976.

BRUNER,J.; SHAPIRO,D.; TAGIURI,R.: The meaning of traits in isolation and in combination.
In: TAGIURI,R.; PETRULLO, L. (Hrsg.): Person perception and interpersonal behavior.
Stanford 1958.

BRUNSWIK, E.: Perception and the representative design of psychological experiments.
Berkeley, 1956.

CATTEL,R.B.; COAN, R.W.: Personality factors in middle childhood as revealled in parents' rating.
Child Developm., 1957, 28, S.439-458.

CATTEL, R.B.; GRUEN, E.: The personality structure of 11-year-old children in terms of behavior rating data.
J.clin.Psychol.,1953,9,S.256-266.

CATTEL,R.B.; PETERSON, D.R.: Personality factors in nursery school children as derived from parent rating.
J.clin.Psychol.,1958,14,S.346-355.

CRONBACH,L.J.;FURBY,L.: How we should measure "change" - or should we?
Psychological Bulletin,1970,74,S.68-80.

DOHSE, W.: Das Schulzeugnis.
Weinheim, 1963.

ELASHOFF,J.; SNOW,R.: Pygmalion auf dem Prüfstand.
München; Kösel, 1972.

FESTINGER,L.: A theory of cognitive dissonance.
Stanford, 1963^3.

FIEDLER,F.E.: A contingency model of leadership.
In: Berkowitz,L.(Hrsg.): Advances in experimental social psychology.
Vol.I, New York, 1964,S.149-190.

FLITNER, A.: Das Schulzeugnis im Lichte neuerer Untersuchungen. Z.f.Päd., 1966, S. 511-558.

GERSTEIN, H.: Erfolg und Versagen im Gymnasium. Beltz, 1972.

GIGERENZER, G.: Artefakte in der dimensionsanalytischen Erfassung von Urteilsstrukturen. Zeitschrift f. Sozialpsychologie,1978,9, S.110-116.

GIGERENZER,G.;STRUBE,G.: Zur Revision der üblichen Anwendung dimensionsanalytischer Verfahren. In: Zeitschrift f. Entwicklungspsychologie und Pädagogische Psychologie, 1978,X,1, S.75-86.

Göller,A.: Zensuren und Zeugnisse. Stuttgart, 1968.

GRAUMANN,C.F.: Eigenschaften als Problem der Persönlichkeitsforschung. In: THOMAE,H.; LERSCH,Ph.(Hrsg.): Hdb.d.Psychol.,IV,Persönlichkeitstheorie und Persönlichkeitsforschung,S.81-155, Göttingen, 1960.

GROOT de,A.D.: Fünfen und Sechsen. Weinheim, 1971.

HANKE,B.;LOHMÖLLER,J.;MANDL,H.: Schülerbeurteilung im Längsschnitt. Universität Augsburg. Forschungsstelle für pädagogische Psychologie und Entwicklungspsychologie.1978.

HANKE,B.;MANDL,H.: Erwartungseffekte in der Schule. Pädagogische Welt,29,1975,S.726-739.

HECKHAUSEN,H.: Die Interaktion der Sozialisationsvariablen in der Genese des Leistungsmotivs, in: GRAUMANN,S.F.(Hrsg.): Hdb.d.Psychol.,Bd.7/II-Sozialpsychologie,1972,S.955-1019.

HECKHAUSEN,H.: Lehrer-Schüler-Interaktion. In: WEINERT,F.E.,et.al. (Hrsg.): Pädagogische Psychologie 1, Funkkolleg, Frankfurt, Fischer, 1974,S.547-573.

HEIDER,F.: Psychologie der interpersonalen Beziehungen. Stuttgart, Klett,1977.

HELSON,H.: Adaptation-level as a basis for a quantitative theory of frame of reference. Ps.R.55,1948.

HUBER,G.; MANDL,H.: Spiegeln Lehrerurteile über Schüler die implizite Persönlichkeitsstruktur der Beurteiler oder der Beurteilungsbögen? unveröff.Manuskript, 1979.

INGENKAMP,K.H.: Die Fragwürdigkeit der Zensurengebung. Weinheim, 1971.

IRLE,M.: Lehrbuch der Sozialpsychologie. Göttingen, 1975.

KEIL,W.: Implizite Persönlichkeitstheorie und Urteile über andere. In: Zeitschrift f.Soz.psychol.,1970,1, S.351-360.

KLEBER,E.W.: Tendenzen, die das Urteil des Lehrers beeinflussen. In: KLEBER,E.W. u.a.: Beurteilung und Beurteilungsprobleme, Weinheim,1976,S.39-61.

KLEBER,E.W.: Die Beurteilerfunktion des Lehrers. In: KLEBER,u.a.: Beurteilung und Beurteilungsprobleme. Weinheim, 1976,S.17-38.

KLEBER,E.W.: Probleme des Lehrerurteils, In: KLAUER,K.J.: Hdb.d.pädagogischen Diagnostik,Band 3; Düsseldorf: Schwann, 1978,S.587-617.

KLEITER,E.F.: Über Theorie und Modell kategorialer Fehler des Lehrerurteils. Psychol.Beiträge,1973,15,S.185-229.

KLEITER,E.F.; PETERMANN,F.: Abbildung von Lernwegen. München, 1977.

KLEITER,E.F.: Über Referenz-Interaktions-und Korrelationsfehler im Lehrerurteil; Taxonomie und Sammelreferat von Kategorienfehlern. In: SCHWARZER,Ch.;SCHWARZER,R.(Hrsg.): Diagnostik im Schulwesen. Braunschweig,1977,S.37-56.

KOGAN,N.: Educational implications of cognitive styles. In: LESSER,G.S.(Hrsg.): Psychology and Educational Practice. 1971,S.242-292.

KRAPP,A.; MANDL,H.: Vorhersage und Erklärung der Schulleistung. Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 1976,8,S.192-219.

KÜHN,W.: Einführung in die multidimensionale Skalierung. München, 1976; UTB 604.

KUTSCHER,J.(Hrsg.): Beurteilen oder verurteilen. München, 1977.

LAUCKEN,U.: Naive Verhaltenstheorie. Stuttgart, Klett, 1974.

LIENERT,G.: Testaufbau und Testanalyse. Weinheim, 1969[3].

LACEY,J.I.; LACEY,B.C.: The law of imitial value in the longitudinal study of autonomic constitution. Ann.New York,Acad.Sci.,1962,98,S.1257-1290.

LOHMÖLLER,J.-B.;MANDL,H.; HANKE,B.: Lehrerspezifische implizite Persönlichkeitstheorien bei der Schülerbeurteilung? Zt.f.Entwicklungspsychol. u.Päd.Psychol., 1976,VIII,S.99-105.

LORD,F.M.: Elementary Models of Measuring Change. In: HARRIS,C.W.: Problems in measuring change. University of Wisconsin Piers,Madison,1969.

MASENDORF,F.; TSCHERNER,K.; TÜCKE,M.: Clusteranalytisch ermittelte Beurteilertendenzen bei der Einschätzung der Schülerpersönlichkeit durch den Lehrer. In: Zeitschrift f.Entwicklungspsychol. und Päd.Psychol., 6, 1974,S.19-27.

MASENDORF,F.; ROEDER,B.: Typologisierung lernschwacher Schüler mit Hilfe der Konfigurationsfrequenzanalyse (KFA). In: Psychologie in Erziehung und Unterricht, 21,1974,S.327-334.

McNEMAR,Q.: On Growth Measurement. Educ.Psychol.Measurement,18,1958,S.47-55.

MERTON,R.K.: The self-fulfilling-prophecy. Ant.Review,8,1948,S.193-210.

MEYER,W.U.; BUTZKAMM,A.: Ursachenerklärung von Rechennoten. In: Lehrerattribuierungen. Zeitschrift f.Entwicklungspsychol.,7,1975, S.53-66.

MITTENECKER,E.: Planung und statistische Auswertung von Experimenten, Wien, 1966[6].

MURRAY,H.A.: Explorations in personality. New York, 1938.

PETERMANN,F.: Veränderungsmessung. Stuttgart, Kohlhammer, 1978.

RICHARDS,J.N.: A simulation study of use of change measure to compare educational programs. American Educational Research Journal, 12,1975,S.299-311.

RICHARDS,J.N.; KARWEIT,N.; PREVATT,T.W.: A FORTRAN program for simulating education and growth with varying school impact. Educational and Psychological Measurement, 35,1975,S.152-156.

ROLLETT,B.; BARTRAM,M.(Hrsg.): Einführung in die hierarchische Clusteranalyse. Stuttgart, 1976.

ROSENBERG,S.; NELSON,G.; VIVEKANANTHAN,P.: A multidimensional approach to the structure of personality impressions. J.o.Personality and Social Psychology, 1968,9.S.283-294.

ROSENTHAL,R.; FODE,K.: The effect of experimenter bias on the performance of the albino rat. Behavioral Science,1963,S.183-189;

ROSENTHAL,R.; FODE,K: Three experiments in experimenter bias. Ps.Rep., 1963,12,S.491-511.

ROSENTHAL,R. et.al.: The role of the reserach assistant in the mediation of experimenter bias. J.Pers., 1963,31,S.313-335.

ROSENTHAL,R.; LAWSON,R.: A longitudinal study of the effects of experimenter bias on the operant learning of laboratory rats J.of.Psychiatric Research,1964,2.S.61-72.

ROSENTHAL,R.: Experimenter outcome-orientation and the results of the psychological experiment. Ps.Bull., 1964,61,S.405-412.

ROSENTHAL,R.et.al.: Experimenters' hypothesis-confirmation and mood as determinants of experimental results. Percept.Mot.Skills,1965,20,S.1237-1252.

ROSENTHAL,R.: Experimenter effects in behavioral research. New York,1966.

ROSENTHAL,R.;FRIEDMAN,N.; KURLAND,D.: Instruction-Reading behavior of the experimenter as an unintended determinant of experimental results. J.Exp.Res. in Personality, 1966,1,S.221-226.

ROSENTHAL,R.; JACOBSON,L.: Teachers' expectancies: Determinants of pupils' IQ-gains. Ps.Rep., 1966,19(1),S.115-118.

ROSENTHAL,R.: Covert communication in the psychological experiment. Ps.Bull.,1967,67,S.356-367.

ROSENTHAL,R.; EVANS,J.: Unpubl.data,Harvard University,1968.

ROSENTHAL,R.; JACOBSON,L.: Pygmalion in the classroom: Teacher expectation and pupils' intellectual development. New York,1968.

ROSENTHAL,R.; ANDERSON,D.: Teacher behavior and the mediation of teacher expectancy effects. Unpubl.data, Harvard University,1969.

ROSENTHAL,R.: Teacher expectations. Persönliche Mitteilung durch Manuskript vom 20. März 1970.

SCHELSKY,H.: Schule und Erziehung in der industriellen Gesellschaft. Würzburg, 1962.

SCHWARZER,Ch.: Lehrerurteil und Schülerpersönlichkeit. München, Kösel, 1976.

SCHWARZER, Ch.; SCHWARZER,R. (Hrsg.): Diagnostik im Schulwesen. Westermann, Braunschweig, 1977.

SCHWARZER,Ch.: Urteile über Kinder in der Eingangsstufe - ein Beitrag zur Dimensionalität und Sozialschichtabhängigkeit des Lehrerurteils. In: SCHWARZER,Ch.; SCHWARZER,R. (Hrsg.): Diagnostik im Schulwesen. Braunschweig,1977,S.57-67.

SCHWARZER,Ch.; SCHWARZER,R.: Praxis der Schülerbeurteilung. München, 1977.

TSCHERNER,K.; MASENDORF,F.: Analyse von Schülerbeurteilungen und Zeugnisnoten. Psychologie in Erziehung und Unterricht. 1974,21,S.135-149.

TUCKER,L.R.: Implications of factor analytis of three-way matrices for measurement of change. In: HARRIS,C.W.(Hrsg.): Problems in measuring change. Madison,1963,S.122-137.

WEISS,R.: Zensur und Zeugnis.
Linz, 1965.

WISHNER,J.: Reanalysis of 'impressions of personality'.
In: Psychol.Rev. 67,1960,S.96-112.

WRIGLEY,C.; NEUHAUS,J.: The matching of two sets of factors.
A.Psychologist,10,1955,S.418-419.

WRIGLEY,C.; NEUHAUS,J.: The matching of two sets of factors.
A.Psychologist,46,1955,S.31-41.

ZIEGENSPECK,J.: Historische Entwicklung und Aussagefälligkeit der tradierten Schülerbeurteilung. Kritische Anmerkungen zur Zensurenproblematik.
In: KUTSCHER,J.(Hrsg.): Beurteilen oder verurteilen.
München,Urban + Schwarzenberg,1977,S.36-58.

ZIELINSKI,W.: Probleme der Beurteilung von Schülerleistungen.
In: Funkkolleg Pädagogische Psychologie (29. Kollegstunde), Tübingen, 1973,S.1 - 15.

ANHANG

1. Instruktion der Versuchspersonen S.281
2. Faktorenmatrix der Stereotypenbeurteilung S.282
3. Faktorenmatrizen der 11 Meßzeitpunkte
   - Eigenwertkriterium- S.283
4. Faktorenmatrizen der 11 Meßzeitpunkte
   - 4-Faktoren-Lösung - S.294
5. Tabelle der ersten zehn Eigenwerte zur
   Abschätzung des Ranges der Korrelations-
   matrizen S.3o5
6. Faktorenmatrix der schiefwinkligen Rotation
   mit Δ = - o,5, Zeitpunkt 5 S.3o6

Instruktion der Vpn

Vielen Dank für Ihre Mitarbeit an dieser Untersuchung. Im nächsten halben Jahr wollen wir versuchen, die Entwicklung Ihrer Erstklässler zu verfolgen.

Dazu ist es notwendig, daß Sie - wie abgemacht - alle 14 Tage je einen Beurteilungsbogen über 20 Schüler einer Ihrer Klassen ausfüllen. Bütte füllen Sie diese Bögen spontan und zügig aus, dann hält sich Ihr Mehraufwand an Arbeit mit ca. 20 Minuten alle 14 Tage in Grenzen.

Bitte'verbessern' Sie auch einmal ausgefüllte Bögen nicht mehr.

Sprechen Sie bitte nicht mit Ihren Kollegen über Ihre Beurteilungen - vielleicht werden Sie dadurch von ihnen beeinflußt.
Da diese Untersuchung weder überwacht werden kann noch soll, bitte ich Sie, diese Instruktionen ernsthaft zu berücksichtigen.
Die Beurteilungsblätter werden also an dem mit Ihnen besprochenen Tag, an dem Sie zur Bearbeitung Zeit haben, in Ihr Fach gelegt und am nächsten Tag wieder abgeholt.

Nochmals vielen Dank für Ihre Hilfe.

(Der Untersuchungsgegenstand selbst war bereits in einer Lehrerkonferenz als ein Tagesordnungspunkt besprochen worden, in der auch die Vpn durch persönliches Ansprechen rekrutiert worden waren).

	Var.	FACTOR 1	FACTOR 2	FACTOR 3	FACTOR 4	FACTOR 5	COMMUNALITY	PCT OF VAR unrot.
1	aufgeweckt	.74318	.11423	.38683	.15503	.13958	.75852	45.9
2	ehrlich	.34247	.61632	.30441	-.13572	.04833	.60320	14.8
3	ehrgeizig	.57465	.43459	.04359	-.00484	.14841	.54360	6.6
4	sprachl.begabt	.29811	.12386	.47875	.10298	-.13677	.35494	4.4
5	gt.Mitarbeit	.77866	.34709	.21971	.14798	.03427	.79812	3.6
6	fleißig	.56226	.69908	.13000	-.16775	.06309	.85388	3.4
7	gt.Lehrerkontakt	.33798	.16348	.67727	-.13731	.31819	.70229	2.5
8	gt.Mitarbeit	.65133	.40527	.38438	.05986	.01554	.74005	2.4
9	hübsch	.03601	-.00385	.65308	.13483	.00433	.44608	2.1
10	höflich	.02131	.76459	.17766	.09989	.31762	.72748	1.9
11	intelligent	.81800	.29776	.16743	.23383	.10236	.85097	1.7
12	kontaktfreudig	.23262	-.10595	.74797	.41711	.25587	.86424	1.6
13	konzentriert	.45098	.78727	.01661	.22299	-.12491	.88879	1.5
14	kreativ	.68060	.19071	.28160	.22824	.18081	.65684	1.3
15	ausdauernd	.53294	.66219	-.05187	-.02657	.01843	.72616	1.1
16	ordentlich	.38289	.74074	.23404	.10277	.12274	.77571	1.1
17	paßt s.Klasse an	.08482	.54464	.41891	.24339	.04374	.54046	.8
18	pos.Einst.Schule	.39937	.68763	.32176	.03220	.16782	.76211	.7
19	akzeleriert	.38080	.02981	.08628	.56559	.26782	.56799	.7
20	sensibel	.06637	.45682	-.28985	-.53814	.07993	.59250	.7
21	gt.häusl.Milieu	.36862	.48193	.18984	.25957	.66760	.91725	.5
22	unterrichtsstör.	-.00207	.79352	-.27707	-.15604	-.07502	.67518	.4
23	unterfordert	.64409	.17260	.13501	.54072	-.06796	.75987	.4
24	begabt	.80694	.25312	.25674	.19838	.02648	.82119	.3
25	selbständig	.51486	.25925	.21140	.50251	-.01202	.62964	.2
26	sympathisch	.21690	.40553	.59729	.03982	-.22626	.62195	.1
27	off.Charakter	.33838	.18276	.60503	.23363	.19456	.60639	-.1
28	vital	.23082	-.12022	.44116	.62321	.05395	.65365	-.2
29	zuverlässig	.44655	.80202	.13904	-.08537	.00992	.86936	-.2

Rotierte Faktorenmatrix "Stereotypen"

MESSZEITPUNKT 1

ROTIERTE FAKTORENMATRIX Eigenwertlösung

	FACTOR 1	FACTOR 2	FACTOR 3	FACTOR 4	FACTOR 5	h^2
1 aufgeweckt	.76531	-.09651	.39122	-.18529	.05857	.78584
2 ehrlich	.38042	.61182	.28838	-.18099	.12781	.65130
3 ehrgeizig	.65149	.40574	.04528	-.03138	-.04383	.59399
4 sprachl.begabt	.28370	.09459	.29727	.15062	.65490	.62938
5 eig.Mitdenken	.80951	.31702	.16929	.14424	.12060	.81357
6 fleißig	.60635	.65775	.08790	-.18752	.10595	.85441
7 gt.Lehrerkontakt	.35500	.19651	.76370	-.00078	.05435	.75083
8 gt.Mitarbeit	.68382	.37469	.30360	.06978	.23206	.75902
9 hübsch	-.00081	.01274	.71226	.13388	.26654	.59645
10 höflich	.06864	.83643	.25112	.06960	-.21860	.82002
11 intelligent	.84099	.26485	.15958	.22368	.06647	.85454
12 kontaktfreudig	.22466	-.02950	.75315	.48861	.05864	.86087
13 konzentriert	.50799	.73236	-.10267	.14516	.16650	.85374
14 kreativ	.73549	.15921	.30970	.22054	-.00953	.71093
15 ausdauernd	.59983	.61457	-.07979	-.09705	.00625	.75332
16 ordentlich	.42568	.75142	.16316	.09253	.13121	.79673
17 paßt s.Klasse an	.03835	.65434	.30776	.32911	.23659	.68863
18 pos.Einst.Schule	.43495	.76575	.26004	.04412	.14357	.77744
19 akzeleriert	.39536	.08389	.10868	.67999	-.22726	.68918
20 sensibel	.13280	.41294	-.17720	-.67982	-.21198	.72664
21 gt.häusl.Milieu	.42589	.54737	.37714	.25450	-.38541	.63646
22 unterrichtsstör.	.02650	.75943	-.37955	-.20751	.05387	.76671
23 unterfordert	.65887	.15081	.00648	.55511	.18767	.80026
24 begabt	.82145	.22111	.19337	.21265	.15628	.83080
25 selbständig	.54284	.25972	.12512	.52733	.07523	.66152
26 sympathisch	.20651	.48045	.42440	.04197	.55431	.69216
27 off.CHarakter	.34626	.22347	.65638	.26994	.04378	.67506
28 vital	.20675	-.06927	.37499	.71914	.12083	.71993
29 zuverlässig	.49703	.76455	.07029	-.12007	.12400	.86631

MESSZEITPUNKT 2

ROTIERTE FAKTORENMATRIX

EIGENWERTLÖSUNG

		FACTOR 1	FACTOR 2	FACTOR 3	FACTOR 4	FACTOR 5	FACTOR 6	h^2
1	aufgeweckt	.74816	-.00030	.48252	.20282	.06324	.08463	.84487
2	ehrlich	.07630	.59915	.11266	.52467	.06080	.04817	.65879
3	ehrgeizig	.43620	.03613	.23699	-.04687	.04356	.17178	.68470
4	sprachl.begabt	.08071	.11090	.08004	.14099	.11484	.92564	.91510
5	eig.Mitdenken	.78180	.23210	.37865	.15130	.05222	-.06658	.83851
6	fleißig	.41666	.77872	.03773	.01565	.11227	.10229	.80476
7	gt.Lehrerkontakt	.29958	.15950	.74865	.25624	.08346	.11862	.76236
8	gt.Mitarbeit	.49206	.35523	.59415	-.07245	.28344	-.01641	.80717
9	hübsch	.17703	-.02323	.09159	.79195	.09186	.09178	.68431
10	höflich	.17876	.73681	.07387	.32361	-.03743	.11944	.69789
11	intelligent	.83984	.31956	.06806	.23019	.05665	.16362	.89817
12	kontaktfreudig	.26267	-.07473	.82995	.04889	-.00962	.02203	.86214
13	konzentrriert	.49974	.77044	.00586	.07577	.11160	-.08798	.86929
14	kreativ	.82752	.08331	.28380	.14767	.10325	-.14910	.82696
15	ausdauernd	.50669	.73788	.08174	-.05118	.04359	.05908	.81589
16	ordentlich	.33024	.67102	.01988	.05559	.02793	-.01461	.87221
17	paßt s.Klasse an	.04635	.33006	.50832	.06318	-.26161	.27571	.51792
18	pos.Einst.Schule	.21504	.72043	.30026	.07700	.12043	.17358	.70597
19	akzeleriert	.63616	.16024	.14737	-.24499	-.47987	.13476	.76056
20	sensibel	.12743	.29151	-.13877	.09022	.74764	.07167	.69272
21	gt.häusl.Milieu	.12935	.38481	.42713	-.11725	.42335	.14110	.56013
22	unterrichtsstör.	.18544	.83153	-.11121	-.01132	.16525	-.05717	.76891
23	unterfordert	.83126	.15362	.05670	-.01396	-.19213	-.02918	.75577
24	begabt	.83085	.30557	.19543	.10666	.11568	.20663	.88933
25	selbständig	.73924	.22622	.12965	.29169	-.20312	.07121	.74587
26	sympathisch	.25646	.30408	.25544	.77216	.04888	.05192	.82480
27	off.Charakter	.17316	-.10715	.80704	.19950	-.22036	-.08245	.78759
28	vital	.35196	-.08567	.36529	-.13201	-.63026	-.09746	.68882
29	zuverlässig	.24483	.81936	-.00746	.10366	.07069	.02741	.74730

MESSZEITPUNKT 3

ROTIERTE FAKTORENMATRIX Eigenwertlösung

		FACTOR 1	FACTOR 2	FACTOR 3	FACTOR 4	FACTOR 5	FACTOR 6	h^2
1	aufgeweckt	.32159	.57600	.56103	.25656	.16053	.03031	.84246
2	ehrlich	.41122	.12722	.25490	.62664	.51991	-.03992	.91483
3	ehrgeizig	.77348	.23633	.28490	-.12554	-.02164	.18496	.78573
4	sprachl.begabt	-.06860	-.66960	.25740	.17376	-.03316	.07386	.55608
5	eig.Mitdenken	.56111	.57303	.38554	.10444	.15627	.07602	.83295
6	fleißig	.80825	.26757	.08888	.20646	.04313	.09729	.78672
7	gt.Lehrerkontakt	.17125	.17049	.78713	.34039	.10414	.00452	.80469
8	gt.Mitarbeit	.56207	.26177	.60496	.01537	.15644	.05175	.77844
9	hübsch	-.03306	.17995	.37355	.53442	.08947	.29403	.55308
10	höflich	.45906	-.01489	.11165	.56399	.16302	-.01581	.56833
11	intelligent	.47345	.73197	.28670	.16430	.07403	-.01387	.87480
12	kontaktfreudig	-.00448	.13155	.86397	.22096	.01084	-.22312	.86249
13	konzentriert	.80389	.34442	.00781	.13122	.14024	-.08466	.80898
14	kreativ	.42931	.61420	.35063	.12598	.18304	.24072	.79181
15	ausdauernd	.74968	.32187	.08686	.14280	.24637	-.13632	.77284
16	ordentlich	.79269	.22046	.06596	.28037	.11870	.18396	.80785
17	paßt s.Klasse an	.32583	-.12700	.26245	.68829	-.00774	-.21546	.71140
18	pos.Einst.Schule	.67852	.18927	.37469	.15947	.28691	.14176	.76446
19	akzeleriert	.10247	.37815	.11759	.08443	.60489	-.27649	.61679
20	sensibel	.30968	-.11326	-.15321	.12270	.10801	.81861	.82904
21	gt.häusl.Milieu	.29120	-.09638	.18139	.12740	.82728	.16391	.85448
22	unterrichtsstör.	.68764	-.05392	-.47449	.32231	-.01206	.02907	.80577
23	unterfordert	.24133	.75191	.10556	.04607	-.06055	-.23143	.69410
24	begabt	.43607	.73713	.21710	.24157	.03478	.00252	.84197
25	selbständig	.29462	.63882	.30592	.29305	.23275	-.23294	.78279
26	sympathisch	.15049	.27215	.16690	.79186	.05861	.16838	.78344
27	off.Chrakter	.11754	-.02448	.78919	.18118	.24855	-.18906	.76758
28	vital	-.03758	.32977	.24039	.03222	.41080	-.61597	.71717
29	zuverlässig	.85205	.22688	.06332	.23918	.11503	.09256	.86048

MESSZEITPUNKT 4

ROTIERTE FAKTORENMATRIX Eigenwertlösung

	FACTOR 1	FACTOR 2	FACTOR 3	FACTOR 4	FACTOR 5	FACTOR 6	FACTOR 7	h^2
1 aufgeweckt	.39584	.54594	.35790	.44948	-.05727	-.11997	-.13897	.82184
2 ehrlich	.44477	.08298	-.03449	.43836	.47251	.39781	.06539	.78386
3 ehrgeizig	.67928	.48691	-.03515	-.10083	-.14861	-.15911	.02362	.75786
4 sprachl.begabt	.03611	.02191	.15230	.10731	-.03423	-.04148	.91676	.87983
5 eig.Mitarbeit	.57322	.43569	.36076	.36887	-.01200	-.03736	-.25083	.84907
6 fleißig	.79862	.16102	.08600	.15885	.16303	.08619	.00819	.73042
7 gt.Lehrerkontakt	.19111	.66555	.21501	.17924	.38599	.14853	-.00963	.72898
8 gt.Mitarbeit	.49773	.64748	.18359	.14699	.16519	-.13847	-.03058	.76967
9 hübsch	.09451	.12695	-.01333	.82221	-.02532	.09228	.15459	.73431
10 höflich	.40369	.09436	-.10964	.08919	.13823	.66953	-.14630	.68063
11 intelligent	.53988	.25428	.44261	.54564	-.02475	-.02966	-.07161	.85637
12 kontaktfreudig	-.04179	.78504	.27596	.18350	-.16864	.20624	.00357	.79885
13 konzentriert	.78501	.04813	.27024	.14416	-.10541	.16786	-.18383	.78545
14 kreativ	.42580	.45435	.31005	.24933	-.01024	.04772	-.32738	.65559
15 ausdauernd	.77537	-.05210	.35388	.10234	.02503	.18863	-.21592	.82245
16 ordentlich	.83774	.01607	.00207	.14668	.14565	.11347	.13766	.77662
17 paßt s.Klasse an	.05984	.19447	.26778	.19406	-.08642	.72417	.04028	.68428
18 pos.Einst.Schule	.67226	.49191	.02736	.09527	.25442	.10243	.13276	.79658
19 akzeleriert	.20746	.12245	.67749	-.00106	.06666	.01781	.05596	.52492
20 sensibel	.19565	-.23422	-.66593	.15184	.18796	-.21306	-.20953	.68428
21 gt.häusl.Milieu	.12949	-.00632	.01335	-.01796	.95486	-.01892	-.04614	.93155
22 unterrichtsstör.	.64541	-.16267	-.37286	-.10482	.03990	.32507	.21936	.74840
23 unterfordert	.31682	.23208	.70543	.24587	.16226	-.07999	.12356	.76031
24 begabt	.56937	.24639	.45033	.47549	.04845	-.12729	-.13320	.85007
25 selbständig	.38895	.10958	.61346	.47213	-.19829	.14092	-.13655	.84035
26 sympathisch	.17284	.22776	-.18155	.72216	.14808	.26741	-.02069	.73010
27 off.Charakter	-.08015	.78020	.28581	.11762	-.04469	.18337	.07927	.75256
28 vital	.01611	.32591	.70534	-.17162	.03745	.03355	-.03936	.63752
29 zuverlässig	.82577	.02527	.15184	.27427	.19599	.06779	.02742	.82457

MESSZEITPUNKT 5

ROTIERTE FAKTORENMATRIX Eigenwertlösung

		FACTOR 1	FACTOR 2	FACTOR 3	FACTOR 4	FACTOR 5	FACTOR 6	h^2
1	aufgeweckt	.82537	.29737	.05550	.20076	.13095	-.02665	.83307
2	ehrlich	.59703	.20818	.19856	.35579	.08369	.12082	.58740
3	ehrgeizig	.64173	.19890	.52655	.01351	-.04568	-.00772	.73307
4	sprachl.begabt	-.01515	-.04753	.31742	.82698	.44169	.13039	.71095
5	eig.Mitdenken	.87728	.20001	.19856	.11105	.07866	-.02298	.86810
6	fleißig	.69501	.10821	.49968	.21519	.05622	-.10629	.60519
7	gt.Lehrerkontakt	.50739	.55243	.06223	.29329	.12320	.18657	.70250
8	gt.Mitarbeit	.64274	.50533	.29167	.11282	.11700	.04001	.78157
9	hübsch	.18182	.21629	.01020	.81093	-.23659	-.05175	.79621
10	höflich	.38482	.17454	.58139	.24173	.26394	-.16471	.67189
11	intelligent	.88757	.02365	.13214	.17410	.16199	.04487	.86436
12	kontaktfreudig	.03807	.83936	.11383	.26586	-.03450	-.02169	.79127
13	konzentriert	.76964	.13577	.32498	.10502	.16740	-.03230	.75649
14	kreativ	.78452	.23324	.23322	-.00393	-.04961	.09741	.73622
15	ausdauernd	.86922	.10534	.16776	.11211	.05848	-.12636	.82673
16	ordentlich	.71356	.02544	.47009	.17476	.05523	-.12106	.77904
17	paßt s.Klasse an	.32794	.54235	.11779	.04002	.45155	.06427	.62520
18	pos.Einst.Schule	.65649	.32337	.46609	.13356	.24236	-.00705	.82942
19	akzeleriert	.59260	.19575	.05169	-.14761	.35297	.04296	.54037
20	sensibel	.16120	-.26884	.77653	-.04263	-.18541	.14543	.75860
21	gt.häusl.Milieu	.37091	.09460	-.09482	.03326	.77036	-.07770	.75611
22	unterrichtsstör.	-.00555	.04555	.02530	.00310	-.02257	.93978	.88646
23	unterfordert	.83378	.09477	.00422	-.11316	.20285	-.18686	.79306
24	begabt	.85750	.01758	.07962	.18520	.22211	.07702	.83152
25	selbständig	.89300	.05960	-.12761	.15847	.06056	.09842	.85575
26	sympathisch	.42192	.27810	-.02927	.60395	.22515	-.04137	.67337
27	off.Charakter	.16037	.86296	.01941	.17373	.10706	.13274	.83006
28	vital	.12081	.73758	-.26293	-.15819	.00158	-.14535	.67391
29	zuverlässig	.63349	.15798	.41993	.11755	.24119	.18001	.70700

MESSZEITPUNKT 6

ROTIERTE FAKTORENMATRIX Eigenwertlösung

		FACTOR 1	FACTOR 2	FACTOR 3	FACTOR 4	FACTOR 5	FACTOR 6	h2
1	aufgeweckt	.69625	.10968	.42840	.08843	.16232	.03030	.71540
2	ehrlich	.21899	.38160	.19980	-.01440	.66505	.21010	.72014
3	ehrgeizig	.32758	.61368	.30817	.38255	-.01314	-.12446	.74069
4	sprachl.begabt	.15918	.01033	-.10333	.86025	.02583	.13193	.79423
5	eig.Mitdenken	.73725	.32953	.28841	.30660	.11091	-.02171	.84208
6	fleißig	.32836	.75575	.16694	.18339	.09562	-.19491	.78839
7	gt.Lehrerkontakt	.31567	.19651	.81289	.05895	.12636	-.17309	.84846
8	gt.Mitarbeit	.47397	.27236	.62644	.17179	.09825	-.12832	.74689
9	hübsch	.13493	.02295	-.00830	.00744	.86213	.02679	.76284
10	höflich	-.04596	.45942	.14568	.13136	.20377	.54224	.58721
11	intelligent	.77610	.29385	.10358	.14618	.20874	-.11284	.77706
12	kontaktfreudig	.26877	-.19278	.85763	.13613	.05288	.10553	.87739
13	konzentriert	.46415	.69544	.09544	-.04855	.11986	-.04338	.72679
14	kreativ	.78752	.12944	.19123	.25895	.09718	.13704	.76879
15	ausdauernd	.41252	.65471	.15256	.05766	.20148	.14017	.68566
16	ordentlich	.16291	.76474	-.12039	.06520	.20477	-.01232	.67220
17	paßt s.Klasse an	-.06203	.27836	.60177	.04700	.10850	.26344	.52684
18	pos.Einst.Schule	.19899	.68414	.31058	.20163	.03774	.18460	.68026
19	akzeleriert	.63720	.12152	.28683	.09840	.00250	.18248	.54605
20	sensibel	-.35926	.23061	-.36149	.28266	.06765	-.56967	.72192
21	gt.häusl.Milieu	.09905	.18696	.18348	.66841	-.11772	-.11931	.55329
22	unterrichtsstör.	-.20054	.74814	-.22457	.09933	.01385	.16067	.68624
23	unterfordert	.79904	.10787	.02197	.29361	.01972	-.00094	.73718
24	begabt	.76218	.34495	.12353	-.17021	.14428	-.17906	.79702
25	selbständig	.84521	.14918	.08749	-.24133	-.00243	.05437	.80549
26	sympathisch	.01601	.28861	.05411	-.20077	.54296	-.39989	.58150
27	off.Charakter	.18097	-.04310	.78998	-.12285	-.07057	.12280	.69383
28	vital	.55695	-.17388	.44694	-.20293	-.23061	.34436	.75313
29	zuverlässig	.23572	.76659	.01683	-.16085	.04785	-.21206	.71671

MESSZEITPUNKT 7

ROTIERTE FAKTORENMATRIX Eigenwertlösung

		FACTOR 1	FACTOR 2	FACTOR 3	FACTOR 4	FACTOR 5	FACTOR 6	FACTOR 7	h2
1	aufgeweckt	.66189	.21403	.45813	-.23395	.07152	.02247	-.08869	.76293
2	ehrlich	.58341	.31327	.51249	.29721	-.21585	.14369	-.13333	.87456
3	ehrgeizig	.29765	.71815	.12710	.13631	.34064	-.01179	-.29619	.77748
4	sprachl.begabt	-.13046	-.11651	.11134	.26318	.06386	.91653	-.01215	.92418
5	eig.Mitdenken	.67819	.43144	.12823	-.05173	.09703	-.10595	-.13661	.70316
6	fleißig	.29472	.75179	.05436	.18994	.10298	-.02475	.11477	.77722
7	gt.Lehrerkontakt	.20962	.36876	.51459	-.15164	.44237	.25309	.02702	.72788
8	gt.Mitarbeit	.29476	.58773	.45726	-.19336	.12113	-.11914	-.02498	.69444
9	hübsch	.04779	-.05085	.01868	-.08947	.00137	-.03172	.03926	.67773
10	höflich	.26886	.34484	.11741	.41879	-.60625	.19402	.18624	.81750
11	intelligent	.82421	.39188	.15244	.01675	.13835	-.17023	.06114	.91104
12	kontaktfreudig	.29381	.08863	.63119	-.04296	.14273	.25340	-.08376	.87121
13	konzentriert	.38528	.21313	.02233	.13527	-.09981	-.11961	.04459	.83429
14	kreativ	.72137	.21657	.19155	-.15944	.10723	-.34918	.04532	.76193
15	ausdauernd	.29793	.05595	.09241	-.17974	-.01195	-.09769	.01862	.87163
16	ordentlich	.41327	.64714	.15035	.24816	-.16242	.18545	.02303	.73526
17	paßt s.Klasse an	.07634	-.00641	.21083	.01659	.03479	-.00227	.00415	.86928
18	pos.Einst.Schule	.49452	.43857	.36054	.17966	.24134	.20122	.01169	.70745
19	akzeleriert	.70621	-.01407	-.09719	-.11797	.02868	.05038	.35625	.65256
20	sensibel	-.07829	.07831	.09222	.87746	.07106	.04211	.00309	.79757
21	gt.häusl.Milieu	.29722	.14294	.09066	.21367	.60143	.05332	.14133	.66092
22	unterrichtsstör.	-.07503	.05797	-.21037	.41247	-.42186	.10244	.06781	.77366
23	unterfordert	.50515	.38619	-.25979	-.26572	.28985	.09224	.21347	.64448
24	begabt	.73356	.42417	.19125	-.05579	.13448	-.07851	.07637	.78433
25	selbständig	.77503	.25963	.08119	-.27914	-.03958	.10153	-.01378	.78487
26	sympathisch	.12359	.21889	.69353	.18787	-.00293	-.11127	.38635	.73643
27	off.Charakter	.12016	.16141	.66943	-.42148	.06161	.38421	.08423	.82389
28	vital	.14057	-.05734	.21622	-.77125	.02852	-.22820	-.02996	.71726
29	zuverlässig	.05238	.09134	.21944	.22664	-.05162	.09437	-.02304	.77437

MESSZEITPUNKT 8

ROTIERTE FAKTORENMATRIX Eigenwertlösung

		FACTOR 1	FACTOR 2	FACTOR 3	FACTOR 4	FACTOR 5	FACTOR 6	FACTOR 7	h^2
1	aufgeweckt	.60219	.21806	.58495	.20752	-.07967	.02567	.15724	.83749
2	ehrlich	.36337	.15693	.16259	.04895	.16078	-.10342	.78821	.87507
3	ehrgeizig	.25233	.81421	.21579	-.11051	.06577	-.15519	-.03189	.82002
4	sprachl.begabt	-.02770	.01705	.09094	.05227	.10052	-.02713	.01392	.94579
5	eig.Mitdenken	.58398	.38954	.43580	.03630	.00818	.15367	.01273	.76017
6	fleißig	.36278	.71514	.13594	-.03371	.20383	-.01995	.10812	.73800
7	gt.Lehrerkontakt	.06128	.19882	.73971	.34009	.06641	.04384	.20945	.75641
8	gt.Mitarbeit	.43861	.43690	.55227	-.04611	-.10313	.13078	.05876	.75662
9	hübsch	.21504	.12104	.25182	.68024	-.14682	-.07800	.14725	.79642
10	höflich	.00638	.31660	-.07422	.07187	.75958	-.06931	.25824	.76745
11	intelligent	.79403	.38606	.18257	.32917	.11111	-.01820	.09542	.94229
12	kontaktfreudig	.33518	-.13915	.68120	.24147	.14686	.23459	.02931	.74184
13	konzentriert	.23339	.72659	-.08637	.27621	.14585	.14743	.04257	.78024
14	kreativ	.76107	.27956	.19583	.00419	-.02711	.00083	.00997	.72321
15	ausdauernd	.10356	.60467	.24481	.14134	.10312	.37580	.36187	.77621
16	ordentlich	.22371	.79186	.11360	.16504	.21652	-.04542	.09569	.79702
17	paßt s.Klasse an	.12759	.23447	.29300	.06166	.77532	-.06859	-.03311	.76784
18	pos.Einst.Schule	.43755	.44782	.23648	.10890	.23752	-.24642	.28790	.69971
19	akzeleriert	.78879	-.18831	-.04687	-.03558	.31847	.29719	.22196	.90141
20	sensibel	-.07822	.02381	-.20618	.08395	.00299	-.82767	.30519	.83622
21	gt.häusl.Milieu	.00006	.14222	.28653	.29474	.09260	.02674	.80860	.86315
22	unterrichtsstör.	-.16553	.61649	-.37941	-.09550	.29871	-.07933	.19597	.74732
23	unterfordert	.62184	.33573	.09173	.20889	-.26596	.30125	-.03836	.72174
24	begabt	.57941	.33085	.11036	.55785	.20690	.12863	.15218	.85216
25	selbständig	.44538	.40078	.12352	.28700	.09186	.44090	.13431	.67885
26	sympathisch	.09919	-.05071	.26258	.76349	.17651	-.16912	.15219	.75453
27	off.Charakter	.06457	.07672	.80513	.06570	.17047	.26129	.24028	.82268
28	vital	.18113	-.09362	.23308	-.17969	-.23556	.73556	.24057	.79695
29	zuverlässig	.50940	.37059	.15925	.19024	.43810	-.02056	.18665	.70244

MESSZEITPUNKT 9 - ROTIERTE FAKTORENMATRIX EIGENWERTLÖSUNG

		FACTOR 1	FACTOR 2	FACTOR 3	FACTOR 4	FACTOR 5	FACTOR 6	FACTOR 7
1	aufgeweckt	.63647	.21056	.22867	.45704	.11828	.00686	-.29199
2	ehrlich	.26484	.66439	.19956	.10115	.15894	.00374	-.01934
3	ehrgeizig	.60376	.46543	.17522	.01671	-.12942	.01431	-.13431
4	sprachl.begabt	-.10972	.28984	.03336	.02236	.11793	.97331	.14372
5	eig.Mitdenken	.86922	.12098	.07821	.17392	.06171	.00446	-.08151
6	fleißig	.57994	.58216	.14292	-.02722	.00775	-.00394	.02380
7	gt.Lehrerkontakt	.02759	.02392	.64801	.13945	-.06612	.23355	-.10976
8	gt.Mitarbeit	.75891	.13476	.26205	.20423	.10063	.06730	-.16581
9	hübsch	.11746	.08016	.82653	.09859	.04561	-.00087	.12914
10	höflich	.28438	.70314	-.12798	.13452	-.06228	.00058	.31199
11	intelligent	.66661	.45838	.20154	.12196	.06955	.00108	-.18748
12	kontaktfreudig	.25724	-.23302	.46992	.44778	.13973	.47761	-.11441
13	konzentriert	.73625	.33579	.03319	-.04649	.07770	-.09243	.21898
14	kreativ	.84502	-.07316	.11157	-.04775	.13188	.02228	.13326
15	ausdauernd	.69112	.32569	.20583	.05344	.17653	-.31216	.10833
16	ordentlich	.63785	.46532	.17952	-.16982	.04734	.10158	.07490
17	paßt s.Klasse an	-.01466	.07303	.10166	-.01874	-.01334	.09662	.86911
18	pos.Einst.Schule	.50041	.55656	.45452	-.11911	.09038	.23372	-.11343
19	akzeleriert	.42091	.16402	.11673	-.30557	.61232	-.32862	-.10316
20	sensibel	-.03628	.07998	.13710	-.34163	-.44246	-.05769	-.14688
21	gt.häusl.Milieu	.11506	.72296	.19796	.11247	-.02198	.11049	-.16028
22	unterrichtsstör.	.14541	.74613	-.20393	-.18616	-.00078	.11965	.26537
23	unterfordert	.56969	.12410	-.06778	.06771	.43594	.12172	-.41275
24	begabt	.58961	.30294	.39825	-.21682	.32657	-.09976	-.05172
25	selbständig	.64283	.33657	.09459	.29953	.17626	-.22881	.08749
26	sympathisch	.27919	.38256	.55350	.18608	.21424	-.23611	.14920
27	off. Charakter	.10221	.07915	.15785	.86202	.10173	.02978	-.00856
28	vital	.09523	-.10619	.12237	.26015	.79344	.22487	-.00683
29	zuverlässig	.59011	.63187	.04744	.18116	-.01905	.03741	-.04755

MESSZEITPUNKT 1o ROTIERTE FAKTORENMATRIX EIGENWERTLÖSUNG

		FACTOR 1	FACTOR 2	FACTOR 3	FACTOR 4	FACTOR 5	FACTOR 6
1	aufgeweckt	.61796	.10885	.46494	.24433	-.04482	-.14390
2	ehrlich	.37265	.49804	.21508	.32156	.01971	.17746
3	ehrgeizig	.69525	.14612	.09307	-.06236	.31923	.19571
4	sprachl.begabt	-.02562	.38636	-.13295	.51953	-.15009	.54521
5	eig.Mitdenken	.75883	.14797	.24608	.23055	-.03133	-.17294
6	fleißig	.53857	.56874	.12952	.03053	.18539	.08143
7	gt.Lehrerkontakt	.40476	-.04054	.64721	.30767	-.01081	-.03668
8	gt.Mitarbeit	.71526	.24705	.37070	.10098	-.04751	-.13068
9	hübsch	.08521	.09930	.36527	.72742	.11870	-.03218
10	höflich	.09647	.58262	-.03162	.21397	.27693	.20302
11	intelligent	.77791	.20729	.12186	.36636	.07068	.11739
12	kontaktfreudig	.18468	-.16209	.73061	.25804	-.26576	.06704
13	konzentriert	.53561	.69118	.05958	.00535	.02061	.00465
14	kreativ	.80782	.05373	.10061	.17551	.11439	-.15490
15	ausdauernd	.73170	.50837	.04767	.06621	-.02119	-.10395
16	ordentlich	.45341	.69360	.02476	.15161	.12966	-.06643
17	paßt s.Klasse an	.02885	.04301	.01398	.02140	.06825	.87422
18	pos.Einst.Schule	.63609	.38212	.14989	.28741	.20709	.08378
19	akzeleriert	.66703	.05199	.11867	-.11228	-.19043	.13442
20	sensibel	.23161	.19312	-.13713	.22582	.74662	.08672
21	gt.häusl.Milieu	.17775	.40990	.55036	-.06517	.44742	.31515
22	unterrichtsstör.	.00526	.82022	-.20261	-.05664	.07953	.00720
23	unterfordert	.64853	.06370	.23963	-.09446	-.30773	.06767
24	begabt	.72177	.25203	.04911	.26425	.14289	.21060
25	selbständig	.58224	.08511	.04576	.48693	-.21539	.01992
26	sympathisch	.26508	.05515	.13488	.71750	.25034	.07998
27	off.Charakter	.15679	-.01407	.82455	.05239	-.15174	-.05208
28	vital	.21143	-.14365	.18538	-.01614	-.70325	.04975
29	zuverlässig	.56020	.40985	.14757	.21080	.08565	.17806

MESSZEITPUNKT II

ROTIERTE FAKTORENMATRIX Eigenwertlösung

		FACTOR 1	FACTOR 2	FACTOR 3	FACTOR 4	FACTOR 5	FACTOR 6	h^2
1	aufgeweckt	.28261	.09287	.22716	.30564	-.07752	.17815	.74269
2	ehrlich	.30464	.30364	.65582	.07759	.14568	.11536	.65576
3	ehrgeizig	.69116	.22356	.09154	.13582	.08707	-.18697	.58867
4	sprachl.begabt	.48441	.07893	.13939	-.13364	-.23605	.40178	.49627
5	eig.Mitdenken	.36342	.53199	.14696	.45459	.13681	-.01722	.66227
6	fleißig	.83703	.19434	.07492	.14931	.14388	.03255	.78665
7	gt.Lehrerkontakt	.13724	.77061	.28940	.06586	.04413	.08170	.70937
8	gt.Mitarbeit	.38174	.56059	.01239	.14515	-.01817	.01572	.69734
9	hübsch	.06153	.07003	.80639	-.01150	-.16857	.04879	.68817
10	höflich	.40974	-.10129	.44384	.07958	.49179	.15936	.64673
11	intelligent	.46246	.35756	.26729	.51447	.04614	.09227	.68887
12	kontaktfreudig	-.07174	.74278	.33145	.10245	-.23311	.08737	.73906
13	konzentriert	.80536	.02459	.03342	.18299	-.02376	-.07044	.68969
14	kreativ	.22647	.53422	.08091	.43618	.09357	-.30518	.62743
15	ausdauernd	.03191	.02829	.05395	.15791	.00294	.72482	.55503
16	ordentlich	.79821	.04023	.16412	.11153	.07394	.02359	.68098
17	paßt s.Klasse an	-.02036	.20795	.00002	.04159	.06913	.75769	.62744
18	pos.Einst.Schule	.68434	.30653	.19838	.13398	.15926	.13381	.66403
19	akzeleriert	.02662	.51997	-.00369	.71691	-.17242	.17644	.57566
20	sensibel	.04674	-.04017	.08797	-.03393	.81343	.04953	.67560
21	gt.häusl.Milieu	.23374	.20671	.58816	-.01434	.16039	-.20491	.51121
22	unterrichtsstör.	.69674	-.37651	.13527	-.01318	.18141	.05441	.65549
23	unterfordert	.47436	.15289	-.00270	.59115	-.09923	.04318	.60954
24	begabt	.54718	.34303	.30276	.49130	-.01558	-.00317	.75038
25	selbständig	.33429	.52377	.06184	.53645	-.02921	.00162	.68142
26	sympathisch	-.00329	.28932	.77382	.16196	.10786	.12242	.73389
27	off.Charakter	-.05237	.67683	.08187	-.06030	-.29536	.31751	.65922
28	vital	-.14772	.21734	.07073	.29139	-.74333	.09007	.71537
29	zuverlässig	.50685	.30835	.25647	.17269	.39116	.14949	.62396

MESSZEITPUNKT 1

ROTIERTE FAKTORENMATRIX 4-Faktoren-Lösung

		FACTOR 1	FACTOR 2	FACTOR 3	FACTOR 4	h^2
1	aufgeweckt	.13537	.71560	.41511	.19593	.74140
2	ehrlich	.61072	.33318	.32192	-.13168	.60494
3	ehrgeizig	.46171	.55088	.07791	.03613	.52418
4	sprachl.begabt	.09050	.30353	.45792	.06428	.31413
5	eig.Mitdenken	.34596	.77458	.23926	.14773	.79875
6	fleißig	.70260	.55618	.15868	-.16584	.85566
7	gt. Lehrerkontakt	.21413	.28643	.69504	.07126	.61606
8	gt. Mitarbeit	.40858	.64203	.40349	.05656	.73867
9	hübsch	-.01945	.02513	.65126	.13373	.44303
10	höflich	.80999	-.06682	.19426	.17373	.72405
11	intelligent	.30571	.91266	.19271	.24538	.85122
12	kontaktfreudig	-.07307	.19443	.76631	.48934	.86066
13	konzentriert	.73867	.48154	.01808	.14319	.79634
14	kreativ	.21135	.66107	.30771	.26421	.64618
15	ausdauernd	.66162	.53572	-.03532	-.03512	.72725
16	ordentlich	.75125	.37448	.24741	.11791	.77972
17	paßt s.Klasse an	.53465	.08712	.40940	.23435	.51597
18	pos.Einst.Schule	.70532	.38192	.34168	.06269	.76401
19	akzeleriert	-.06810	.66431	.09419	.64669	.86443
20	sensibel	.47532	.05161	-.25844	-.50036	.54574
21	gt.häusl.Milieu	.54632	.33275	.25363	.36806	.60899
22	unterrichtsstör.	.73064	.02337	-.27802	-.20426	.65341
23	unterfordert	.14883	.66045	.13293	.48667	.71286
24	begabt	.24722	.80698	.27757	.19935	.82560
25	selbständig	.24411	.52537	.20633	.47221	.60116
26	sympathisch	.34588	.24321	.55478	-.02452	.48828
27	offener Charakter	.20642	.30776	.62256	.26524	.60628
28	vital	-.12240	.22861	.42743	.62317	.63828
29	zuverlässig	.79435	.44890	.15368	-.10148	.86642

MESSZEITPUNKT 2

ROTIERTE FAKTORENMATRIX 4-Faktoren-Lösung

		FACTOR 1	FACTOR 2	FACTOR 3	FACTOR 4	h^2
1	aufgeweckt	.03909	.73127	.39428	.35252	.81601
2	ehrlich	.55175	.09311	.04068	.45935	.51756
3	ehrgiezig	.64771	.41845	.16600	.08364	.63571
4	sprachl. begabt	.19580	.06431	.05897	.28811	.12896
5	eig.Mitdenken	.24637	.77397	.29260	.24817	.80229
6	fleißig	.79110	.39951	-.01732	.07854	.79191
7	gt.Lehrerkontakt	.21310	.29021	.59467	.44228	.68239
8	gt. Mitarbeit	.40816	.47552	.36524	.21936	.57423
9	hübsch	.00916	.15159	.06166	.58855	.37326
10	höflich	.67535	-.13197	.04663	.25533	.54088
11	intelligent	.32858	.83152	.03646	.29431	.88734
12	kontaktfreudig	-.04059	.26341	.90429	.09182	.89721
13	konzentriert	.76278	.49124	-.06472	.08987	.83541
14	kreativ	.09704	.80634	.18685	.23274	.74868
15	ausdauernd	.74261	.30602	.03078	.00382	.80308
16	ordentlich	.56623	.32413	-.00453	.03907	.85703
17	paßt s.Klasse an	.30418	.11023	.45160	.11593	.32116
18	pos.Einst.Schule	.72928	.21348	.19214	.21906	.66233
19	akzeleriert	.14600	.52222	.33722	-.31485	.62132
20	sensibel	.34511	.07643	-.37305	.33214	.37651
21	gt.häusl.Milieu	.41798	.14833	.14471	.20631	.26022
22	unterrichtsstör.	.80115	-.17766	-.18141	.00869	.70640
23	unterfordert	.13995	.81564	.12789	-.06150	.69687
24	begabt	.33714	.82093	.10017	.24878	.86119
25	selbständig	.21601	.71435	.29328	.19906	.62371
26	sympathisch	.29469	.24248	.18610	.68665	.65170
27	off.Charakter	-.08296	.16434	.78091	.23551	.70615
28	vital	-.11298	.36210	.55626	-.28843	.53650
29	zuverlässig	.79622	.23804	-.04294	.11394	.70546

MESSZEITPUNKT 3

ROTIERTE FAKTORENMATRIX 4-Faktoren-Lösung

		FACTOR 1	FACTOR 2	FACTOR 3	FACTOR 4	h^2
1	aufgeweckt	.63957	.21044	.56674	.21008	.81867
2	ehrlich	.22213	.87261	.14578	.23891	.98743
3	ehrgeizig	.71276	.22166	.15422	-.25834	.64768
4	sprachl.begabt	-.45785	.16135	.22595	-.26573	.35368
5	eig.Mitdenken	.79560	.25091	.33634	.09624	.81786
6	fließig	.69078	.49165	.04211	-.15456	.74455
7	gt.Lehrerkontakt	.25657	.25197	.80475	.08399	.78239
8	gt.Mitarbeit	.58691	.26609	.49658	-.01747	.66212
9	hübsch	.11907	.27773	.43105	-.01528	.27738
10	höflich	.21734	.61641	.19181	-.03190	.46747
11	intelligent	.84666	.17518	.25428	.24009	.85970
12	kontaktfreudig	.12274	.05240	.84612	.24667	.79457
13	konzentriert	.71971	.47993	-.06354	.04421	.75437
14	kreativ	.73507	.20715	.33698	.04337	.69868
15	ausdauernd	.66386	.49453	.03236	.11238	.69899
16	ordentlich	.64692	.57134	.05477	-.21983	.79592
17	paßt s.Klasse an	.07002	.55216	.33139	.04143	.42132
18	pos.Einst.Schule	.57603	.49328	.32748	-.10282	.69296
19	akzeleriert	.29228	.22264	.16525	.43047	.34760
20	sensibel	.16433	.27427	-.07155	-.64422	.52237
21	gt.häusl.Milieu	.12498	.51134	.22276	.01604	.32748
22	unterrichtsstör.	.30769	.58689	-.42973	-.22648	.67497
23	unterfordert	.65251	-.03451	.16899	.39116	.58474
24	begabt	.79850	.20672	.20710	.23125	.77426
25	selbständig	.59601	.27734	.31065	.47505	.75433
26	sympathisch	.22579	.51559	.31601	.04656	.40863
27	off.Charakter	.08178	.23812	.73243	.19411	.63753
28	vital	.15111	.09357	.21131	.68785	.54937
29	zuverlässig	.68866	.57714	.02459	-.17342	.83802

MESSZEITPUNKT 4

ROTIERTE FAKTORENMATRIX 4-Faktoren-Lösung

		FACTOR 1	FACTOR 2	FACTOR 3	FACTOR 4	h^2
1	aufgeweckt	.57547	.10081	.53735	.26722	.71914
2	ehrlich	.19368	.55689	.07037	.53001	.63350
3	ehrgeizig	.36772	.48693	.25842	-.09729	.44856
4	sprachl.begabt	.01377	.00029	.10966	.03860	.01371
5	eig.Mitdenken	.67974	.30162	.41565	.22490	.77637
6	fleißig	.51049	.04759	.07210	.13837	.70432
7	gt. Lehrerkontakt	.19970	.32807	.65005	.19441	.61408
8	gt. Mitarbeit	.38649	.42357	.55126	.05400	.66199
9	hübsch	.19577	.04298	.13922	.66064	.49600
10	höflich	.14970	.43461	.01638	.29862	.30087
11	intelligent	.74458	.19108	.30903	.36691	.82138
12	kontaktfreudig	.15862	-.03410	.77880	.19167	.66908
13	konzentriert	.71601	.42559	.01103	.09596	.70136
14	kreativ	.51932	.23965	.41677	.16823	.52912
15	ausdauernd	.74120	.42727	-.03857	.05788	.73678
16	ordentlich	.47143	.03070	-.08190	.14653	.71377
17	paßt s.Klasse an	.19952	.04631	.27096	.29692	.20353
18	pos.Einst.Schule	.30953	.71815	.37987	.11699	.76955
19	akzeleriert	.47096	-.01538	.33050	-.08370	.33829
20	sensibel	-.22650	.32583	-.44500	.14689	.37705
21	gt.häusl.Milieu	-.06503	.41729	.03555	.03015	.18053
22	unterrichtsstör.	.07207	.01390	-.32836	.08774	.49759
23	unterfordert	.61279	.06465	.43941	.05114	.57539
24	begabt	.74856	.23569	.29712	.26331	.77351
25	selbständig	.80897	-.07989	.24647	.33074	.83093
26	sympathisch	.07540	.26590	.15397	.75013	.66279
27	off.Charakter	.09352	-.00443	.79040	.13174	.65086
28	vital	.37604	-.13294	.59621	-.21789	.46280
29	zuverlässig	.00816	.02156	-.02638	.21162	.80168

MESSZEITPUNKT 5

ROTIERTE FAKTORENMATRIX 4-Faktoren-Lösung

		FACTOR 1	FACTOR 2	FACTOR 3	FACTOR 4	h^2
1	aufgeweckt	.82595	.30915	.20006	-.01342	.81552
2	ehrlich	.58527	.19774	.38164	.13305	.54500
3	ehrgeizig	.65999	.16576	.08815	.45940	.68189
4	sprachl.begabt	.11877	-.00119	.54973	.12625	.33225
5	eig.Mitdenken	.88625	.19732	.11919	.13457	.85669
6	fleißig	.72989	.06105	.27270	.39939	.77318
7	gt.Lehrerkontakt	.50064	.53511	.32318	.03701	.64280
8	gt.Mitarbeit	.66967	.49548	.16457	.23241	.77506
9	hübsch	.13207	.19018	.60048	.02273	.41470
10	höflich	.50068	.12970	.32320	.34083	.48813
11	intelligent	.89771	.02503	.20247	.04451	.84949
12	kontaktfreudig	.06061	.76777	.27268	.07898	.67374
13	konzentriert	.80557	.11532	.18153	.20682	.73798
14	kreativ	.75373	.21788	.04118	.21268	.66251
15	ausdauernd	.86462	.09387	.12701	.09860	.78223
16	ordentlich	.74659	-.00214	.22792	.35836	.73777
17	paßt s.Klasse an	.42517	.47949	.15911	-.01650	.43627
18	pos.Einst.Schule	.72698	.31471	.22610	.35416	.80410
19	akzeleriert	.61837	.21246	-.05352	-.00231	.43039
20	sensibel	.18809	-.27517	.04918	.72163	.63427
21	gt.häusl.Milieu	.47042	.14026	.11958	-.19836	.29462
22	unterrichtsstör.	-.04002	.06117	.08447	.08389	.01952
23	unterfordert	.86139	.09469	-.11184	-.09137	.77182
24	begabt	.86879	.02694	.22196	-.01498	.80500
25	selbständig	.84299	.07794	.13662	-.15609	.75974
26	sympathisch	.43547	.26321	.57485	-.12329	.60457
27	off.Charakter	.16745	.88231	.20253	.01600	.84779
28	vital	.11697	.61938	-.12277	-.23048	.46551
29	zuverlässig	.68715	.13862	.24258	.28505	.63149

MESSZEITPUNKT 6

ROTIERTE FAKTORENMATRIX 4-Faktoren-Lösung

		FACTOR 1	FACTOR 2	FACTOR 3	FACTOR 4	h^2
1	aufgeweckt	.64927	.19632	.[illegible]902	.11864	.67539
2	ehrlich	.21190	.55625	.22906	-.03685	.40165
3	ehrgeizig	.28328	.55678	.22844	.49264	.68504
4	sprachl.begabt	.13367	-.34874	-.08136	.75443	.59603
5	eig.Mitdenken	.70390	.35154	.28772	.37476	.84048
6	fleißig	.29262	.74489	.08345	.30361	.73956
7	gt.Lehrerkontakt	.28340	.29260	.71369	.11013	.69456
8	gt.Mitarbeit	.43063	.33[illegible]	.56045	.23443	.66295
9	hübsch	.13955	.26565	.03291	-.09436	.11197
10	höflich	-.02727	.46[illegible]	.19611	.13437	.21752
11	intelligent	.74527	.36644	.10838	.17183	.73102
12	kontaktfreudig	.22092	-.12162	.89023	.14354	.87647
13	konzentriert	.43487	.69786	.07368	.05444	.68451
14	kreativ	.74162	.15335	.25244	.26659	.71938
15	ausdauernd	.37160	.66558	.16926	.14124	.62935
16	ordentlich	.13579	.74486	-.12539	.15851	.61584
17	paßt s.Klasse an	-.03927	.27673	.52190	.09302	.35926
18	pos.Einst.Schule	.15876	.61642	.29363	.31218	.58884
19	akzeleriert	.56930	.13625	.33837	.13671	.47618
20	sensibel	-.29882	.20244	-.48934	.26159	.41086
21	gt.häusl.Milieu	.10815	.11782	.09597	.55753	.32206
22	unterrichtsstör.	-.19366	.60816	-.21097	.18998	.49184
23	unterfordert	.75172	.10754	.06610	.31791	.68201
24	begabt	.74987	.42227	.11139	-.11872	.76711
25	selbständig	.81786	.17505	.15565	-.16543	.75114
26	sympathisch	.05746	.43681	-.04491	-.17844	.22796
27	off.Charakter	.15434	-.09756	.75260	-.05467	.59327
28	vital	.49325	-.19714	.54625	-.09836	.58371
29	zuverlässig	.21211	.72946	-.04244	-.01139	.57394

MESSZEITPUNKT 7

ROTIERTE FAKTORENMATRIX 4-Faktoren-Lösung

		FACTOR 1	FACTOR 2	FACTOR 3	FACTOR 4	h^2
1	aufgeweckt	.60365	.25997	.49113	-.14607	.69453
2	ehrlich	.44843	.30873	.48653	.43191	.71965
3	ehrgeizig	.31311	.68950	.15426	.04579	.59934
4	sprachl.begabt	-.26562	-.15876	.32624	.33289	.31301
5	eig.Mitdenken	.64548	.45102	.14103	-.00584	.63999
6	fleißig	.33852	.70296	.07604	.28940	.69828
7	gt.Lehrerkontakt	.21929	.34583	.60663	-.08539	.54298
8	gt.Mitarbeit	.33371	.56984	.40245	-.06714	.60256
9	hübsch	.04624	-.01454	.64239	.00474	.41503
10	höflich	.17113	.24102	.09940	.61987	.48149
11	intelligent	.88167	.36744	.15102	.05632	.93833
12	kontaktfreudig	.23299	.06065	.79361	-.41777	.86231
13	konzentriert	.41919	.75915	-.00130	.18262	.78538
14	kreativ	.73572	.25349	.15170	-.17154	.65798
15	ausdauernd	.32018	.86965	.09122	-.04486	.86914
16	ordentlich	.36868	.55551	.19047	.43877	.67331
17	paßt s.Klasse an	.16614	-.08608	.25502	.09030	.10820
18	pos.Einst.Schule	.44090	.40189	.44627	.24586	.61551
19	akzeleriert	.61029	.01933	.07938	-.03613	.38044
20	sensibel	-.09778	.00447	-.00507	.72647	.53737
21	gt.häusl.Milieu	.31658	.13118	.19148	.06359	.15814
22	unterrichtsstör.	-.08214	.46067	-.23485	.57599	.60589
23	unterfordert	.51509	.28896	-.03107	-.16177	.37595
24	begabt	.75074	.40251	.22090	.06120	.77817
25	selbständig	.69939	.32563	.20903	-.12913	.65555
26	sympathisch	.19233	.15099	.54295	.22581	.40557
27	off.Charakter	.03486	.19323	.81655	-.22935	.75791
28	vital	.18097	.03459	.20212	-.69380	.55615
29	zuverlässig	.59130	.44787	.24573	.35661	.73778

MESSZEITPUNKT 8

ROTIERTE FAKTORENMATRIX 4-Faktoren-Lösung

		FACTOR 1	FACTOR 2	FACTOR 3	FACTOR 4	h^2
1	aufgeweckt	.67970	.10076	.51464	.17839	.76883
2	ehrlich	.25797	.33135	.52771	-.10845	.46658
3	ehrgeizig	.35346	.68910	-.00488	.07627	.60563
4	sprachl.begabt	.00852	.02777	.25554	-.23265	.12027
5	eig.Mitdenken	.63452	.28504	.19820	.36174	.65400
6	fließig	.41372	.69116	.12127	.02458	.66419
7	gt.Lehrerkontakt	.21830	.15301	.68277	.22594	.58847
8	gt.Mitarbeit	.52852	.27774	.30975	.30268	.54586
9	hübsch	.38508	-.00175	.52000	-.20811	.46200
10	höflich	-.04890	.59522	.25616	-.22880	.47465
11	intelligent	.90218	.31396	.27882	-.11169	1.00272
12	kontaktfreudig	.37995	-.13015	.60178	.29162	.60849
13	konzentriert	.37697	.64012	-.01981	.09467	.56120
14	kreativ	.71722	.20919	.06413	.13990	.58185
15	ausdauernd	.23997	.57619	.27709	.40242	.62830
16	ordentlich	.35244	.75268	.17640	-.05347	.72472
17	paßt s.Klasse an	.10926	.44347	.32987	-.06452	.32158
18	pos.Einst.Schule	.42429	.50018	.38662	-.18669	.61453
19	akzeleriert	.52997	.03175	.16082	.11162	.32020
20	sensibel	-.17556	.12625	.06680	-.63244	.45121
21	gt.häusl.Milieu	.02971	.27653	.63542	.05884	.48458
22	unterrichtsstör-	-.13377	.67861	-.13085	-.24939	.55773
23	unterfordert	.70976	.12470	.03582	.23749	.57699
24	begabt	.67024	.32510	.37139	-.03780	.69427
25	selbständig	.53623	.34125	.21442	.29253	.53554
26	sympathisch	.21140	.00756	.54417	-.20862	.38438
27	off.Charakter	.12421	.11052	.70122	.45378	.72527
28	vital	.18510	-.14631	.08228	.69941	.55161
29	zuverlässig	.48549	.48529	.35566	-.10332	.60837

MESSZEITPUNKT 9

ROTIERTE FAKTORENMATRIX 4-Faktoren-Lösung

		FACTOR 1	FACTOR 2	FACTOR 3	FACTOR 4	h^2
1	aufgeweckt	.60349	.20872	.60989	-.03519	.78097
2	ehrlich	.35749	.54648	.13563	.15723	.46956
3	ehrgeizig	.53990	.48508	.14745	.10283	.55911
4	sprachl.begabt	-.26862	.35289	.28145	.20106	.31633
5	eig.Mitdenken	.76720	.18133	.34351	-.02996	.74037
6	fleißig	.59392	.55540	.02225	.14419	.68249
7	gt.Lehrerkontakt	.41115	.10130	.44710	.50924	.63852
8	gt.Mitarbeit	.67956	.16710	.42838	.14300	.69369
9	hübsch	.19411	.06339	.24113	.57579	.43138
10	höflich	.24301	.69817	-.01881	-.02555	.54750
11	intelligent	.66711	.43863	.28338	.06414	.72185
12	kontaktfreudig	.06905	-.14008	.78635	.38581	.79158
13	konzentriert	.68565	.36266	.01045	.07313	.60710
14	kreativ	.71586	.01639	.16436	.12001	.55414
15	ausdauernd	.78654	.25608	.04244	.12793	.70240
16	ordentlich	.58704	.48328	.02730	.23920	.63613
17	paßt s.Klasse an	.08664	.15595	-.12492	.20852	.09091
18	pos.Einst.Schule	.49543	.53901	.16846	.47546	.79756
19	akzeleriert	.68116	-.03728	-.12629	.10040	.49140
20	sensibel	.00673	.51570	-.29766	.08899	.36251
21	gt.häusl.Milieu	.17442	.61503	.13619	.17991	.45960
22	unterrichtsstör.	.12997	.72493	-.22201	.00546	.59174
23	unterfordert	.59066	.04733	.32843	-.10685	.47040
24	begabt	.72533	.18186	.01476	.36989	.69622
25	selbständig	.67452	.29554	.24110	-.01918	.60068
26	sympathisch	.43871	.26738	.18093	.34861	.41823
27	off.Charakter	.07652	.06705	.61622	-.00455	.39010
28	vital	.20658	-.20301	.39628	.11258	.25360
29	zuverlässig	.53468	.66174	.23843	-.01156	.78076

MESSZEITPUNKT 1o

ROTIERTE FAKTORENMATRIX - 4-Faktoren-Lösung

		FACTOR 1	FACTOR 2	FACTOR 3	FACTOR 4	h^2
1	aufgeweckt	.62407	.50467	-.01353	-.08550	.65165
2	ehrlich	.50204	.23380	.44286	.07230	.50807
3	ehrgeizig	.63492	.10523	.11872	.15576	.45256
4	sprachl.begabt	.04986	.10462	.65424	.04532	.44351
5	eig.Mitdenken	.76348	.36010	-.05641	-.01999	.71616
6	fleißig	.69250	.00619	.35284	.12190	.61895
7	gt.Lehrerkontakt	.37718	.66606	.02334	-.10161	.59676
8	gt.Mitarbeit	.76744	.33056	.01199	-.10417	.70924
9	hübsch	.14077	.58736	.26124	.19738	.47201
10	höflich	.28456	-.04753	.49557	.29699	.41704
11	intelligent	.75360	.36509	.18947	.12833	.75356
12	kontaktfreudig	.11955	.72295	.03382	-.36902	.67427
13	konzentriert	.75003	-.11194	.38653	-.00267	.72449
14	kreativ	.76097	.27706	-.15001	.13830	.69748
15	ausdauernd	.87501	-.00312	.17332	-.01232	.79585
16	ordentlich	.67274	-.05049	.38525	.16886	.63207
17	paßt s.Klasse an	-.02614	.05067	.40648	.03544	.16973
18	pos.Einst.Schule	.70137	.25208	.29351	.21928	.68969
19	akzeleriert	.56380	.13451	-.01581	-.20055	.37643
20	sensibel	.26607	.01856	.21122	.70823	.61734
21	gt.häusl.Milieu	.32670	.21398	.41643	.15362	.34953
22	unterrichtsstör.	.30469	-.40062	.47105	.15211	.49836
23	unterfordert	.57713	.20044	-.02771	-.33890	.48888
24	begabt	.69973	.23412	.25676	.16023	.63604
25	selbständig	.51593	.36267	.11608	-.04282	.41303
26	sympathisch	.24822	.53833	.26168	.34461	.53865
27	off.Charakter	.19646	.55768	-.01202	-.29883	.43905
28	vital	.11428	.20161	-.12833	-.52948	.35052
29	zuverlässig	.62618	.18768	.34917	.08252	.55606

MESSZEITPUNKT 11

ROTIERTE FAKTORENMATRIX 4-Faktoren-Lösung

		FACTOR 1	FACTOR 2	FACTOR 3	FACTOR 4	h^2
1	aufgeweckt	.50313	.52986	.37821	.19165	.71367
2	ehrlich	.31777	.65373	-.16852	.18700	.59172
3	ehrgeizig	.68686	.09160	-.11301	-.08854	.50078
4	sprachl.begabt	.31251	.08602	-.01127	.36039	.23507
5	eig.Mitdenken	.61097	.41815	.18154	-.02668	.58180
6	fleißig	.82325	.11645	-.22182	.13120	.75771
7	gt.Lehrerkontakt	.30221	.63774	.26505	.08195	.57501
8	gt.Mitarbeit	.67201	.28851	.16766	.07269	.56822
9	hübsch	.04305	.52231	-.06021	.14294	.29872
10	höflich	.32911	.33445	-.54943	.18007	.55447
11	intelligent	.66673	.38116	.12194	.11984	.61904
12	kontaktfreudig	.12338	.63041	.50163	.09704	.67369
13	konzentriert	.75182	-.02340	-.17037	.05449	.59778
14	kreativ	.50021	.36784	.24003	-.29236	.52861
15	ausdauernd	.05818	.08182	.08089	.49804	.26467
16	ordentlich	.70958	.09373	-.27020	.14584	.60656
17	paßt s.Klasse an	.01312	.15172	.10442	.53056	.31559
18	pos.Einst.Schule	.68372	.29694	-.14995	.20164	.61878
19	akzeleriert	.27231	.03251	.30063	.10841	.17734
20	sensibel	.04176	.16190	-.49403	-.04581	.27412
21	gt.häusl.Milieu	.23415	.50608	-.20495	-.08514	.36020
22	unterrichtsstör.	.48315	-.10417	-.52074	.16755	.54353
23	unterfordert	.65372	.05323	.18633	.07719	.47087
24	begabt	.73522	.37891	.10918	.05629	.69922
25	selbständig	.62062	.33520	.33884	-.00322	.61235
26	sympathisch	.08733	.74968	-.05298	.15916	.59778
27	off.Charakter	.05821	.37230	.49252	.29945	.47424
28	vital	.01201	.04571	.65447	.13804	.44962
29	zuverlässig	.54996	.40218	-.23835	.15579	.54529

MESSZEITPUNKTE

1		2		3		4		5		6		7		8		9		10		11	
29 F	4 F	29 F	4 F	29 F	4 F	29 F	4 F	29 F	4 F	29 F	4 F	29 F	4 F	29 F	4 F	29 F	4 F	29 F	4 F	29 F	4 F
13,3	13,01	11,62	11,33	12,61	12,32	11,10	10,77	13,87	13,60	10,31	9,98	10,82	10,55	10,55	10,18	11,42	11,06	11,02	10,64	9,55	9,14
4,21	3,97	4,44	4,14	3,43	3,12	3,65	3,22	2,77	2,44	4,01	3,61	3,75	3,40	3,36	2,92	3,18	2,74	3,42	2,96	3,49	3,04
1,91	1,54	2,16	1,78	2,59	2,16	1,92	1,51	1,83	1,38	2,05	1,67	2,67	2,16	2,42	1,93	2,06	1,58	2,01	1,53	2,20	1,68
1,26	,89	1,16	1,36	1,45	1,09	1,56	1,16	1,29	,85	1,44	1,15	1,46	1,05	1,63	1,25	1,42	,90	1,37	,85	1,58	,94
1,05		1,27		1,27		1,43		1,15		1,33		1,34		1,45		1,34		1,22		1,24	
,97		1,03		1,01		1,31		1,07		1,20		1,25		1,38		1,22		1,07		1,06	
,73		,91		,92		1,23		,96		,99		1,09		1,13		1,03		,96		,93	
,69		,72		,71		,81		,76		,96		,92		1,02		1,00		,78		,83	
,62		,65		,65		,71		,67		,80		,82		,84		,85		,76		,75	
,54		,58		,63		,63		,58		,76		,73		,63		,68		,69		,67	

Tabelle der ersten 1o Eigenwerte zur Rangschätzung der Matrizen
(für 29 und 4 extrahierte Faktoren)

		FACTOR 1	FACTOR 2	FACTOR 3	FACTOR 4	COMMUNALITY
1	aufgeweckt	.81606	.15119	-.15165	.11504	.81552
2	ehrlich	.58398	.04560	.04206	.29468	.54500
3	ehrgeizig	.76538	.04803	.53447	-.09792	.68189
4	sprachl.begabt	.08885	-.11441	.12260	.52592	.33225
5	eig.Mitdenken	.91887	.04502	-.03134	-.00658	.85669
6	fleißig	.80212	-.07881	.25976	.10619	.77318
7	gt.Lehrerkontakt	.48484	.41071	-.00563	.24913	.64280
8	gt.Mitarbeit	.71772	.36992	.14180	.02087	.77596
9	hübsch	.07464	.06981	.03739	.59356	.41470
10	höflich	.55083	-.00321	.25714	.19298	.48813
11	intelligent	.90003	-.14398	-.13591	.11049	.84949
12	kontaktfreudig	.05420	.71985	.14944	.21645	.67374
13	konzentriert	.84564	-.03742	.05225	.05167	.73798
14	kreativ	.81050	.09792	.07297	-.09555	.66251
15	ausdauernd	.88729	-.05724	-.07168	.01815	.78223
16	ordentlich	.81421	-.15727	.20580	.07299	.73777
17	paßt s.Klasse an	.41375	.39443	-.05471	.10297	.43627
18	pos.Einst.Schule	.79552	.26647	.23575	.05753	.80410
19	akzeleriert	.63439	.13161	-.11582	-.12560	.43039
20	sensibel	.34658	-.32408	.05134	-.13648	.63427
21	gt.häusl.Milieu	.42100	.05368	-.27838	.12164	.29462
22	unterrichtsstör.	-.03065	.05235	.10096	.08682	.01952
23	unterfordert	.86741	-.01272	-.26934	-.13065	.77182
24	begabt	.85512	-.14029	-.18775	.14850	.80500
25	selbständig	.80679	-.06847	-.32085	.09584	.75974
26	sympathisch	.35326	.10458	-.16599	.57109	.60457
27	off.Charakter	.15695	.83250	.07224	.14353	.84779
28	vital	.08417	.63284	-.19976	-.11211	.46551
29	zuverlässig	.73660	-.00758	.15971	.10522	.63149

Faktorenmatrix "aktuelle Beurteilungen", Zeitpunkt 5
- schiefwinklige Rotation mit = -o,5

Erklärung gemäß § 5,1; 6/7 der Promotionsordnung

Diese Dissertation wurde von mir selbständig verfaßt; ich habe nur die angegebenen Quellen und Hilfsmittel benutzt und wörtlich oder inhaltlich übernommene Stellen als solche gekennzeichnet.

Ferner habe ich diese Dissertation noch nicht als Prüfungsarbeit bei einer akademischen und staatlichen Prüfung verwendet oder mit dieser oder einer anderen Dissertation bereits einen Promotionsversuch unternommen.

Rainer Baumann

Lebenslauf und wissenschaftlicher Werdegang

20.10.1944	Geboren in Bad Landeck/Schlesien Vater : Werner Baumann, Verlagsleiter Mutter: Hildegard Baumann, Hausfrau
1945	Übersiedelung nach Aalen/Württ.
1951 - 1955	Besuch der Volksschule in Aalen
1955 - 1962	Besuch des Schubart-Gymnasiums in Aalen
1962 - 1965	Besuch der Wirtschaftsoberschule in Schwäbisch-Gmünd
SS 1965	Studium generale zur Berufsfindung an der Eberhard-Karls-Universität Tübingen
8.10.1965	Ergänzungsprüfung zum Vollabitur
WS 1965	Einschreibung für das Fach Psychologie an der Eberhard-Karls-Universität Tübingen
22.2.1968	Diplom-Vorprüfung
4.10.1968	Großes Latinum
6.11.1970	Diplom-Hauptprüfung
11.1.1971 - 30.5.1971	Praktikant beim Psychologischen Dienst des Arbeitsamtes Aalen
1971 - 1972	Promotionsstipendium des Bundes, später des Landes Baden-Württemberg; Planung und erste Datensammlung für die Dissertation
1973 - 1975	Wissenschaftliche Begleitung des Vorschulprojektes an der Pädagogischen Hochschule Schwäbisch-Gmünd
1975 - 1979	Wissenschaftlicher Assistent (Vw) am Seminar für Psychologie der Pädagogischen Hochschule Aachen
ab 1980	Klinischer Psychologe am Bethlehem-Krankenhaus Stolberg/Rhld.
Familienstand	verheiratet, 1 Kind
Veröffentlichungen	1. Versuchsplanung und Datenerhebung im Forschungsprojekt Schwäbisch Gmünd. in: Retter,H. (Hrsg.): Schlüsselbegriffe in der Vorschulerziehung, Herder, Freiburg, 1973, S. 199 - 211 2. Untersuchungen zum sozialen und emotional-affektiven Bereich. in: Baumann,R., Diener,K., Retter,H.: Auswirkungen vorschulischer Förderung im ersten Schuljahr, Klinkhardt, Bad Heilbrunn, 1977, S. 187 - 288